

ON THE EQUIVALENCE OF THE ENTROPIC CURVATURE-DIMENSION CONDITION AND BOCHNER'S INEQUALITY ON METRIC MEASURE SPACES

MATTHIAS ERBAR * KAZUMASA KUWADA * KARL-THEODOR STURM

CONTENTS

1. Introduction	1
2. (K, N) -convex functions and their EVI gradient flows	4
2.1. Gradient flows and (K, N) -convexity in a smooth setting	4
2.2. Evolution Variational Inequalities in metric spaces	7
2.3. $\text{EVI}_{K,N}$ implies (K, N) -convexity	10
3. Entropic and Riemannian curvature-dimension conditions	12
3.1. The entropic curvature-dimension condition	12
3.2. Calculus and heat flow on metric measure spaces	17
3.3. The Riemannian curvature-dimension condition	18
4. Equivalence of $\text{CD}^e(K, N)$ and the Bochner Inequality $\text{BI}(K, N)$	22
4.1. From $\text{CD}^e(K, N)$ to Bakry-Émery-Wang $\text{BEW}(K, N)$ and $\text{BI}(K, N)$	22
4.2. From $\text{BEW}(0, N)$ to $\text{EVI}(0, N)$	26
4.3. From $\text{BEW}(K, N)$ to $\text{CD}^e(K, N)$	30
References	34

1. INTRODUCTION

Bochner's inequality is one of the most fundamental estimates in geometric analysis on Riemannian manifolds. It states that

$$\frac{1}{2}\Delta|\nabla u|^2 - \langle \nabla u, \nabla \Delta u \rangle \geq K \cdot |\nabla u|^2 + \frac{1}{N} \cdot |\Delta u|^2 \quad (1.1)$$

for each smooth function u on a Riemannian manifold (M, g) provided $K \in \mathbb{R}$ is a lower bound for the Ricci curvature on M and $N \in (0, \infty]$ is an upper bound for the dimension of M . The main results of this paper is an analogous Bochner inequality on metric measure spaces (X, d, m) with linear heat flow and satisfying the (reduced) curvature-dimension condition. Indeed, we will also prove the converse: if the heat flow on a mms (X, d, m) is linear then an appropriate version of (1.1) (for the canonical gradient and Laplacian on X) will imply the reduced curvature-dimension condition. Besides that, we also derive new, sharp W_2 -contraction results for the heat flow as well as L^2 -gradient estimates and prove that each of them is equivalent to the curvature-dimension condition.

That way, we obtain a complete one-to-one correspondence between the Eulerian picture captured in the Bochner inequality and the Lagrangian interpretation captured in the curvature-dimension inequality.

To simplify the presentation, in this introduction all mms under consideration will be assumed to be non-branching length spaces of unit total mass.

The **curvature-dimension condition** $\text{CD}(K, N)$ was introduced by the third named author in [24]. It was later adopted and slightly modified by Lott & Villani, see also the elaborate presentation in the monograph [25]. The $\text{CD}(K, N)$ -condition for finite N is a sophisticated

tightening up of the much simpler $\text{CD}(K, \infty)$ -condition introduced as a synthetic Ricci bound for metric measure spaces independently by Sturm [24] and Lott & Villani [18]. From the very beginning, a disadvantage of the $\text{CD}(K, N)$ -condition for finite N was the lack of a local-to-global result. To by-pass this drawback, Bacher & Sturm [8] introduced the *reduced curvature-dimension condition* $\text{CD}^*(K, N)$ which shares a local-to-global property and which is equivalent to the local version of $\text{CD}(K, N)$.

The curvature-dimension condition $\text{CD}(K, N)$ has been verified for Riemannian manifolds [24], Finsler spaces [19], Alexandrov spaces [21], cones [7] and warped products of Riemannian manifolds [15]. Actually, in all these cases the conditions $\text{CD}(K, N)$ and $\text{CD}^*(K, N)$ turned out to be equivalent.

A completely different approach to generalized curvature-dimension bounds was set forth in the pioneering work of Bakry and Émery [9]. It applies to the general setting of Dirichlet forms and the associated Markov semigroups and is formulated using the (iterated) *carré du champ* operators build from the generator of the semigroup. This **energetic curvature-dimension condition** $\text{BE}(K, N)$ has proven a powerful tool in particular in infinite dimensional situations. It yields hypercontractivity of the semigroup and has successfully been used to derive functional inequalities like the logarithmic Sobolev inequalities in a variety of examples.

Among the remarkable analytic consequences of the Bakry–Émery condition $\text{BE}(K, \infty)$ we single out the point-wise gradient estimates for the semigroup H_t . It implies that for any f in a large class of functions

$$\Gamma(H_t f) \leq e^{-2Kt} H_t \Gamma(f), \quad (1.2)$$

where Γ is the carré du champ operator.

The relation between the two notions of curvature bounds based on optimal transport and Dirichlet forms has been studied in large generality by Ambrosio, Gigli and Savaré in a series of recent works [5, 6], see also [2].

The key tool of their analysis is a powerful calculus on metric measure spaces which allows them to match the two settings. Starting from a metric measure structure they introduce the so called Cheeger energy which takes over the role of the ‘standard’ Dirichlet energy and is obtained by relaxing the L^2 -norm of the slope of Lipschitz functions. A key result is the identification of the L^2 -gradient flow of the Cheeger energy with the Wasserstein gradient flow of the entropy. This is the mms equivalent of the famous result by Jordan–Kinderlehrer–Otto [14] and allows one to define unambiguously a heat flow in metric measure spaces.

We say that a metric measure space is *Riemannian* if the heat flow is linear. This is equivalent to the Cheeger energy being the associated Dirichlet form. We denote its domain by $W^{1,2}$. Under the assumption of linearity of the heat flow, Ambrosio–Gigli–Savaré prove that $\text{BE}(K, \infty)$ is equivalent to $\text{CD}(K, \infty)$.

Combining linearity of the heat flow with the $\text{CD}(K, \infty)$ condition leads to the **Riemannian curvature condition** $\text{RCD}(K, \infty)$ introduced in [5]. This concept again turns out to be stable under Gromov–Hausdorff convergence and tensorization.

Recently, also **Bochner’s inequality** has been extended to singular spaces. Ohta & Sturm [20] proved it for Finsler spaces and Zhang & Zhu [27] for Alexandrov spaces (extending previous work by Gigli, Kuwada & Ohta [13]). Finally, Ambrosio, Gigli & Savaré established the Bochner inequality without the dimension term (i.e. with $N = \infty$) in $\text{RCD}(K, \infty)$ spaces.

However, in the classical setting, the full strength of Bochner’s inequality only comes to play if also the dimension effect is taken into account, i.e. with finite N . This can be seen for example from the famous results of Li–Yau [17] who derive from it a differential Harnack inequality, eigenvalue estimates for the Laplacian and Gaussian heat kernel bounds.

In this work we extend the results of Ambrosio–Gigli–Savaré and prove for Riemannian metric measure spaces the equivalence of curvature-dimension bounds via optimal transport and via the Bakry–Émery approach. In particular, we establish the full Bochner inequality on such metric measure spaces.

Our approach strongly relies on properties and consequences of a new curvature-dimension condition, the so-called **entropic curvature dimension condition** $\text{CD}^e(K, N)$. It simply states that the Boltzmann entropy Ent is (K, N) -convex on the Wasserstein space $\mathcal{P}_2(X, d)$. Here a function u on an interval $I \subset \mathbb{R}$ is called (K, N) -convex if

$$u'' \geq K + \frac{1}{N} \cdot (u')^2. \quad (1.3)$$

A function U on a geodesic space is called (K, N) -convex if it is (K, N) -convex along each unit speed geodesic. This way, (K, N) -convexity is a weak formulation of

$$\text{Hess } U \geq K + \frac{1}{N} (\nabla U \otimes \nabla U). \quad (1.4)$$

Our first result is the following

Theorem 1 (Thm. 2.10). *A non-branching mms (X, d, m) satisfies the entropic curvature-dimension condition $\text{CD}^e(K, N)$ if and only if it satisfies the reduced curvature-dimension condition $\text{CD}^*(K, N)$.*

We say that a metric measure space satisfies the **Riemannian curvature-dimension condition** $\text{RCD}(K, N)$ if it satisfies $\text{CD}^e(K, N)$ and the heat flow is linear. This notion turns out to have the natural stability properties. Namely, we prove (see Theorems 3.19, 3.20, 3.21) that the $\text{RCD}(K, N)$ condition is preserved under Gromov–Hausdorff convergence and tensorization of metric measure spaces and holds globally if and only if it holds locally.

The geometric intuition coming from the analysis of (K, N) -convex functions and their gradient flows leads to a new form of the **Evolution Variation Inequality** $\text{EVI}_{K, N}$ on the Wasserstein space taking into account also the effect of the dimension bound.

Until now, the notion of $\text{EVI}_{K, N}$ gradient flow was known only without dimension term (i.e. with $N = \infty$). These Evolution Variational Inequalities first appeared in the setting of Hilbert spaces where they characterize uniquely the gradient flows of K -convex functionals.

In a general metric setting and in connection with optimal transport these inequalities have been extensively studied in [11, 5]. In particular, it turned out that $\text{RCD}(K, \infty)$ spaces can be characterized by the fact that the heat flow is an $\text{EVI}_{K, \infty}$ gradient flow of the entropy.

Here we obtain a reinforcement of this result. Namely, the new Riemannian curvature-dimension condition $\text{RCD}(K, N)$ is equivalent to the existence of an $\text{EVI}_{K, N}$ gradient flow of the entropy in the following sense.

Theorem 2 (Def. 1.7, Thm. 2.15). *A mms (X, d, m) satisfies $\text{RCD}(K, N)$ if and only if every $\mu_0 \in \mathcal{P}_2(X, d)$ is the starting point of a curve $(\mu_t)_{t \geq 0}$ in $\mathcal{P}_2(X, d)$ such that for any other $\nu \in \mathcal{P}_2(X, d)$ and a.e. $t > 0$:*

$$\frac{d}{dt} s_{K/N} \left(\frac{1}{2} W_2(\mu_t, \nu) \right)^2 + K \cdot s_{K/N} \left(\frac{1}{2} W_2(\mu_t, \nu) \right)^2 \leq \frac{N}{2} \left(1 - \frac{U_N(\nu)}{U_N(\mu_t)} \right) \quad (1.5)$$

Here $U_N(\mu) = \exp \left(-\frac{1}{N} \text{Ent}(\mu) \right)$ and $s_{K/N}(r) = \sqrt{\frac{N}{K}} \sin \left(\sqrt{\frac{K}{N}} r \right)$ (provided $K > 0$ and with the usual re-interpretation in the case $K \leq 0$).

This curve is unique, in fact, it is the heat flow which we denote with denote in the following with slight abuse of notation by $\mu_t = H_t \mu_0$.

The Evolution Variation Inequality $\text{EVI}_{K, N}$ as stated above immediately implies new, sharp contraction estimates (or, more precisely, expansion bounds) in Wasserstein metric for the heat flow.

Theorem 3 (Thm 1.12, Thm. 3.1). *Let (X, d, m) be a $\text{RCD}(K, N)$ space. Then for any $\mu, \nu \in \mathcal{P}_2(X, d)$ and $s, t > 0$:*

$$s_{K/N} \left(\frac{1}{2} W_2(H_t \mu, H_s \nu) \right)^2 \leq e^{-K(s+t)} s_{K/N} \left(\frac{1}{2} W_2(\mu, \nu) \right)^2 + \frac{N}{K} \left(1 - e^{-K(s+t)} \right) \frac{(\sqrt{t} - \sqrt{s})^2}{2(s+t)}. \quad (1.6)$$

Due to the work of the second name author [16], it is well known that W_2 -expansion bounds are intimately related to L^2 -gradient estimates. The next result is a particular case of a more general equivalence that will be the subject of a forthcoming publication.

Theorem 4 (Thm. 4.4). *Assume that the heat flow is linear and satisfies the W_2 -expansion bound (1.6). Then for any f of finite Cheeger energy:*

$$|\nabla \mathsf{H}_t f|_w^2 + \frac{4Kt^2}{N(e^{2Kt} - 1)} |\Delta \mathsf{H}_t f|^2 \leq e^{-2Kt} \mathsf{H}_t (|\nabla f|_w^2). \quad (1.7)$$

Here $|\nabla f|_w$ denotes the weak upper gradient of f introduced in [4]. Wang [26] has pointed out the relation between L^2 -gradient estimates for the heat flow and the Bochner inequality on smooth Riemannian manifolds. Extending his equivalence results to the non-smooth setting leads to the Bochner formula for the canonical gradients and Laplacians on mms.

Theorem 5 (Thm. 4.5). *Assume that the heat flow is linear and satisfies the L^2 -gradient estimate (1.7). Then for all $f \in D(\Delta)$ with $\Delta f \in W^{1,2}(X, d, m)$ and all $g \in D(\Delta)$ bounded and non-negative with $\Delta g \in L^\infty(X, m)$ we have*

$$\frac{1}{2} \int \Delta g |\nabla f|_w^2 \, dm - \int g \langle \nabla(\Delta f), \nabla f \rangle \, dm \geq K \int g |\nabla f|_w^2 \, dm + \frac{1}{N} \int g (\Delta f)^2 \, dm. \quad (1.8)$$

Theorem 6 (Thm. 4.10). *Assume that the heat flow is linear. Then the Bochner inequality $\text{BI}(K, N)$ (1.8) implies the entropic curvature-dimension condition $\text{CD}^e(K, N)$.*

Thus we have closed the circle. All the previous key properties are equivalent to each other, at least if we require the heat flow to be linear (in other words, if the underlying mms is a Riemannian mms).

Theorem 7 (Summary). *For a Riemannian metric measure space (X, d, m) the following properties are equivalent:*

- (i) $\text{CD}^*(K, N)$,
- (ii) $\text{CD}^e(K, N)$,
- (iii) Existence of the $\text{EVI}_{K,N}$ gradient flow of the entropy starting from every $\mu \in \mathcal{P}_2(X, d)$,
- (iv) The W_2 -expansion bound (1.6),
- (v) The Bakry–Émery–Wang L^2 -gradient estimate (1.7),
- (vi) The Bochner inequality $\text{BI}(K, N)$ (1.8).

Organization of the article. First we illustrate the new concept of (K, N) -convexity in a smooth and finite dimensional setting. Since many of the arguments which relate geodesic convexity, the Evolution Variational Inequality and space-time expansion bounds for the gradient flow are of a purely metric nature we study $\text{EVI}_{K,N}$ and its consequences in the general setting of metric spaces in Section 2. In Section 3 we turn to the study of (K, N) -convexity of the entropy on the Wasserstein space. The entropic curvature-dimension condition is introduced in Section 3.1 and its basic properties are established. In particular we prove equivalence with the reduced curvature-dimension condition for non-branching spaces. In Section 3.3 we prove that the entropic curvature-dimension condition plus linearity of the heat flow is equivalent to the existence of an $\text{EVI}_{K,N}$ gradient flow of the entropy which leads to the Riemannian curvature-dimension condition. Here we also prove the stability results for $\text{RCD}(K, N)$. Finally, in Section 4 we prove the equivalence of the entropic curvature-dimension condition, L^2 -gradient estimates and the Bochner inequality for Riemannian metric measure spaces.

2. (K, N) -CONVEX FUNCTIONS AND THEIR EVI GRADIENT FLOWS

2.1. Gradient flows and (K, N) -convexity in a smooth setting. In order to illustrate the concept of (K, N) -convexity of the entropy and the consequences for its gradient flow, we consider in this section a smooth and finite-dimensional setting.

Let M be a smooth connected and geodesically complete Riemannian manifold with metric tensor $\langle \cdot, \cdot \rangle$ and Riemannian distance d . Let $U : M \rightarrow \mathbb{R}$ be a smooth function. Given two real numbers $K \in \mathbb{R}$ and $N > 0$, we say that U is (K, N) -convex, if

$$\text{Hess } U - \frac{1}{N}(\nabla U \otimes \nabla U) \geq K, \quad (2.1)$$

in the sense that for all $x \in M$ and $v \in T_x M$ we have

$$\text{Hess } U(x)[v] - \frac{1}{N} \langle \nabla U(x), v \rangle_x^2 \geq K |v|_x^2.$$

Obviously, this condition becomes weaker as N increases and in the limit $N \rightarrow \infty$ we recover the notion of K -convexity, i.e. $\text{Hess } U \geq K$. It turns out to be useful to introduce the function $U_N : M \rightarrow \mathbb{R}_+$ given by

$$U_N(x) = \exp \left(-\frac{1}{N} U(x) \right).$$

A direct calculation shows that (2.1) can equivalently be written as:

$$\text{Hess } U_N \leq -\frac{K}{N} \cdot U_N. \quad (2.2)$$

This condition can be thought of as a “concavity” property of U_N . As with concavity, it can be expressed in an integrated form. To this end we introduce the following functions.

Definition 2.1. For $\kappa \in \mathbb{R}$ and $\theta \geq 0$ we define the functions

$$s_\kappa(\theta) = \begin{cases} \frac{1}{\sqrt{\kappa}} \sin(\sqrt{\kappa}\theta), & \kappa > 0, \\ \theta, & \kappa = 0, \\ \frac{1}{\sqrt{-\kappa}} \sinh(\sqrt{-\kappa}\theta), & \kappa < 0, \end{cases}$$

$$c_\kappa(\theta) = \begin{cases} \cos(\sqrt{\kappa}\theta), & \kappa \geq 0, \\ \cosh(\sqrt{-\kappa}\theta), & \kappa < 0. \end{cases}$$

Moreover, for $t \in [0, 1]$ we set

$$\sigma_\kappa^{(t)}(\theta) = \begin{cases} \frac{s_\kappa(t\theta)}{s_\kappa(\theta)}, & \kappa\theta^2 \neq 0 \text{ and } \kappa\theta^2 < \pi^2, \\ t, & \kappa\theta^2 = 0, \\ +\infty, & \kappa\theta^2 \geq \pi^2. \end{cases}$$

Lemma 2.2. The following statements are equivalent:

- (i) The function U is (K, N) -convex.
- (ii) For each constant speed geodesic $(\gamma_t)_{t \in [0, 1]}$ in M and all $t \in [0, 1]$ we have with $d := d(\gamma_0, \gamma_1)$:

$$U_N(\gamma_t) \geq \sigma_{K/N}^{(1-t)}(d) \cdot U_N(\gamma_0) + \sigma_{K/N}^{(t)}(d) \cdot U_N(\gamma_1). \quad (2.3)$$

- (iii) For each constant speed geodesic $(\gamma_t)_{t \in [0, 1]}$ in M we have that

$$U_N(\gamma_1) \leq c_{K/N}(d) \cdot U_N(\gamma_0) + \frac{s_{K/N}(d)}{d} \cdot \frac{d}{dt} \Big|_{t=0} U_N(\gamma_t). \quad (2.4)$$

Proof. (i) \Rightarrow (ii): Let $(\gamma_t)_{t \in [0, 1]}$ be a constant speed geodesic. Then in particular $|\dot{\gamma}_t|_{\gamma_t} = d$ and (2.2) immediately yields that the function $u : t \mapsto U_N(\gamma_t)$ satisfies

$$u''(t) \leq -\frac{K}{N} d^2 \cdot u(t). \quad (2.5)$$

The function $v : [0, 1] \rightarrow \mathbb{R}$ given by the right-hand side of (2.3) has the same boundary values as u and satisfies $v'' = -(K/N)d \cdot v$. A comparison principle thus yields $u \geq v$.

(ii) \Rightarrow (iii): This follows immediately by subtracting $U_N(\gamma_0)$ on both sides of (2.3), dividing by t and letting $t \searrow 0$.

(iii) \Rightarrow (i): Let $\gamma : [-1, 1] \rightarrow M$ be a constant speed geodesic with $\gamma_0 = x$ and $\dot{\gamma}_0 = v$, i.e. $d = d(\gamma_0, \gamma_1) = |v|$. Using (2.4) for the rescaled geodesics $\gamma' : [0, 1] \rightarrow M$, $t \mapsto \gamma_{\varepsilon t}$ and $\gamma'' : [0, 1] \rightarrow M$, $t \mapsto \gamma_{-\varepsilon t}$ and adding up we obtain

$$U_N(\gamma_\varepsilon) + U_N(\gamma_{-\varepsilon}) - 2c_{K/N}(\varepsilon d) \cdot U_N(\gamma_0) \leq 0.$$

Dividing by ε^2 and using the fact that $c_{K/N}(\varepsilon d) = 1 - \frac{K}{N}\varepsilon^2 d^2 + o(\varepsilon^2)$ finally yields

$$\text{Hess } U_N(x)[v] \leq -\frac{K}{N}|v|^2.$$

□

Remark 2.3. We note that the existence of a (K, N) -convex function $U : M \rightarrow \mathbb{R}$ with $K > 0$ poses strong constraints on the manifold M . In particular, it implies that the diameter of M is bounded by $\sqrt{\frac{N}{K}}\pi$. This is immediate from the characterization (2.3) and the singularity of the coefficient $\sigma_\kappa^{(t)}(\cdot)$ at $\pi/\sqrt{\kappa}$.

Lemma 2.4. *A smooth curve $x : [0, \infty) \rightarrow M$ is a solution to the gradient flow equation*

$$\dot{x}_t = -\nabla U(x_t) \quad \forall t > 0, \quad (2.6)$$

if and only if the following Evolution Variation Inequality (EVI $_{K,N}$) holds: for all $z \in M$ and all $t > 0$:

$$\frac{d}{dt} s_{K/N} \left(\frac{1}{2} d(x_t, z) \right)^2 + K \cdot s_{K/N} \left(\frac{1}{2} d(x_t, z) \right)^2 \leq \frac{N}{2} \left(1 - \frac{U_N(z)}{U_N(x_t)} \right). \quad (2.7)$$

Proof. To prove the only if part, fix $t \geq 0$, $z \in M$ and a constant speed geodesic $\gamma : [0, 1] \rightarrow M$ connecting x_t to z . Observe that by (2.6) and the first variation formula we have

$$\begin{aligned} \left. \frac{d}{ds} \right|_{s=0} U_N(\gamma_s) &= -\frac{1}{N} U_N(x_t) \langle \nabla U(x_t), \dot{\gamma}_0 \rangle = \frac{1}{N} U_N(x_t) \langle \dot{x}_t, \dot{\gamma}_0 \rangle \\ &= -\frac{1}{N} U_N(x_t) \frac{d}{dt} \frac{1}{2} d(x_t, z)^2. \end{aligned}$$

Combining this with the (K, N) -convexity condition in the form (2.4) we obtain with $d = d(x_t, z)$:

$$U_N(z) \leq c_{K/N}(d) U_N(x_t) - \frac{s_{K/N}(d)}{Nd} U_N(x_t) \frac{d}{dt} \frac{1}{2} d(x_t, z)^2. \quad (2.8)$$

Using the identity

$$\frac{2}{N} s_{K/N} \left(\frac{1}{2} \theta \right)^2 = \frac{1}{K} (1 - c_{K/N}(\theta)),$$

it is immediate to see that the last inequality is equivalent to (2.7).

For the if part, fix $t \geq 0$ and a constant speed geodesic $\gamma : [0, 1] \rightarrow M$ with $\gamma_0 = x_t$. Using the Evolution Variational inequality in the form (2.8) with $z = \gamma_\varepsilon$ for some $\varepsilon > 0$ we obtain

$$U_N(\gamma_\varepsilon) - c_{K/N}(\varepsilon|v|) U_N(\gamma_0) \leq \frac{s_{K/N}(\varepsilon|v|)}{N\varepsilon|v|} U_N(\gamma_0) \langle \dot{x}_t, \varepsilon v \rangle.$$

Dividing by ε and letting $\varepsilon \searrow 0$, taking into account that $c_{K/N}(\varepsilon d) = 1 + o(\varepsilon)$ and $s_{K/N}(\varepsilon d) = \varepsilon d + o(\varepsilon^2)$, we obtain

$$\langle -\nabla U(x_t), v \rangle \leq \langle \dot{x}_t, v \rangle.$$

Since the direction of $v \in T_{x_t}M$ was arbitrary we obtain (2.6). □

We conclude this section by exhibiting some 1-dimensional models of (K, N) -convex functions.

Example 2.5. Each of the following are (K, N) -convex functions. Note that the domain of definition is maximal in each case.

(i) For $N > 0$ and $K > 0$ let $U : (0, \sqrt{N/K}\pi) \rightarrow \mathbb{R}$ defined by

$$U(x) = -N \log(\sin(x\sqrt{K/N})) .$$

(ii) For $N > 0$ and $K \leq 0$ let $U : (0, \infty) \rightarrow \mathbb{R}$ defined by

$$U(x) = -N \log \sinh(x\sqrt{-K/N}) .$$

(iii) For $N > 0$ and $K < 0$ let $U : (-\infty, \infty) \rightarrow \mathbb{R}$ defined by

$$U(x) = -N \log \cosh(x\sqrt{-K/N}) .$$

2.2. Evolution Variational Inequalities in metric spaces. In this section we study the Evolution Variational Inequality with parameters K and N and the associated notion of gradient flow in a purely metric setting. In particular, we investigate the relation with geodesic convexity. Our approach extends the results obtained in [11, 5] where the case $N = \infty$ has been considered.

Let (X, d) be a complete and separable geodesic metric space and let $U : X \rightarrow (-\infty, \infty]$ be a lower semi-continuous functional. We denote by $D(U) := \{x \in X : U(x) < \infty\}$ the proper domain of U . We define the *descending slope* of U at $x \in D(U)$ as

$$|\nabla^- U|(x) := \limsup_{y \rightarrow x} \frac{(U(x) - U(y))_+}{d(x, y)} .$$

For $x \notin D(U)$ we set $|\nabla^- U| = +\infty$.

A curve $\gamma : I \rightarrow X$ defined on an interval $I \subset \mathbb{R}$ is called *absolutely continuous* if

$$d(\gamma_s, \gamma_t) \leq \int_s^t g(r) \, dr \quad \forall s, t \in I, s \leq t, \quad (2.9)$$

for some $g \in L^1(I)$. For an absolutely continuous curve γ the *metric speed*, defined by

$$|\dot{\gamma}|(t) := \lim_{h \rightarrow 0} \frac{d(\gamma_{t+h}, \gamma_t)}{|h|},$$

exists for a.e. $t \in I$ and is the minimal g in (2.9) (see e.g. [3, Theorem 1.1.2]). The following is a classical notion of gradient flow in a metric space, see e.g. [3].

Definition 2.6 (Gradient flow). *We say that a locally absolutely continuous curve $x : [0, \infty) \rightarrow X$ with $x_0 \in D(U)$ is a (downward) gradient flow of U starting in x_0 if the Energy Dissipation Equality holds:*

$$U(x_t) = U(x_s) + \frac{1}{2} \int_s^t |\dot{x}_r|^2 + |\nabla^- U|(x_r) \, dr \quad \forall 0 \leq s \leq t. \quad (2.10)$$

We introduce here a more restrictive notion of gradient flow based on the Evolution Variational Inequality. Given a number $N \in (0, \infty)$ we define the functional $U_N : X \rightarrow [0, \infty)$ by setting

$$U_N(x) := \exp\left(-\frac{1}{N}U(x)\right). \quad (2.11)$$

Definition 2.7 (EVI $_{K,N}$ gradient flow). *Let $K, N \in \mathbb{R}$ with $N > 0$ and let $x : (0, \infty) \rightarrow D(U)$ be a locally absolutely continuous curve. We say that (x_t) is an EVI $_{K,N}$ gradient flow of U starting in x_0 if $\lim_{t \rightarrow 0} x_t = x_0$ and if for all $z \in D(U)$ the Evolution Variational Inequality*

$$\frac{d}{dt} s_{K/N}\left(\frac{1}{2}d(x_t, z)\right)^2 + K \cdot s_{K/N}\left(\frac{1}{2}d(x_t, z)\right)^2 \leq \frac{N}{2} \left(1 - \frac{U_N(z)}{U_N(x_t)}\right) \quad (2.12)$$

holds for a.e. $t > 0$.

Lemma 2.8. *If $(x_t)_t$ is an $\text{EVI}_{K,N}$ flow for U , then it is also an $\text{EVI}_{K',N'}$ flow for U for any $K' \leq K$ and $N' \geq N$. Moreover, (x_t) is an EVI_K flow for U , i.e. for all $z \in D(U)$ and a.e. $t > 0$:*

$$\frac{1}{2} \frac{d}{dt} d(x_t, z)^2 + \frac{K}{2} d(x_t, z)^2 \leq U(z) - U(x_t). \quad (2.13)$$

Proof. Using the identity

$$\frac{2}{N} s_{K/N} \left(\frac{1}{2} \theta \right)^2 = \frac{1}{K} (1 - c_{K/N}(\theta)), \quad (2.14)$$

and one checks that (2.12) is equivalent to either of the following inequalities:

$$\frac{1}{2} \frac{d}{dt} d(x_t, z)^2 \leq \frac{Nd}{s_{K/N}(d)} \left[c_{K/N}(d) - \frac{U_N(z)}{U_N(x_t)} \right] \quad (2.15)$$

$$\frac{1}{2} \frac{d}{dt} d(x_t, z)^2 \leq \frac{d}{s_{K/N}(d)} N \left[1 - \frac{U_N(z)}{U_N(x_t)} \right] - 2Kd \frac{s_{K/N}(\frac{1}{2}d)^2}{s_{K/N}(d)}, \quad (2.16)$$

where we set $d = d(x_t, z)$. Consistency in K can be seen from (2.15) by noting that for any $\theta \geq 0$ we have that $s_{K/N}(\theta)$ is increasing in K and $c_{K/N}(\theta)/s_{K/N}(\theta)$ is decreasing in K . Consistency in N follows from (2.16) and the fact that for any $v \in \mathbb{R}$ and $\theta \geq 0$ both

$$N \left[1 - \exp \left(-\frac{1}{N} v \right) \right] \frac{1}{s_{K/N}(\theta)} \quad \text{and} \quad -K \cdot \frac{s_{K/N}(\frac{1}{2}\theta)^2}{s_{K/N}(\theta)}$$

are increasing in N .

(2.13) follows immediately from (2.16) by passing to the limit as $N \rightarrow \infty$. For this we note that

$$\lim_{N \rightarrow \infty} s_{K/N}(d) = d, \quad \lim_{N \rightarrow \infty} N \left[1 - \frac{U_N(z)}{U_N(x_t)} \right] = U(z) - U(x_t).$$

□

Remark 2.9. This shows consistency with the theory of EVI_K gradient flows of geodesically K -convex functions. It can be thought of as the limiting case $N = \infty$. By taking the limit $N \rightarrow \infty$ in the estimates obtained in this section we recover the corresponding results for EVI_K flows established in [11, 5].

We summarize here some properties of $\text{EVI}_{K,N}$ gradient flows. To this end we set for $\kappa \in \mathbb{R}$ and $t \geq 0$:

$$e_\kappa(t) = \int_0^t e^{\kappa s} ds.$$

Proposition 2.10. *Let (x_t) be an $\text{EVI}_{K,N}$ gradient flow of U starting in x_0 . Then the following statements hold:*

- (i) *If $x_0 \in D(U)$ then (x_t) is also a metric gradient flow in the sense of Definition 2.6. In particular, the map $t \mapsto U(x_t)$ is non-increasing.*
- (ii) *We have the uniform regularization bound*

$$\frac{U_N(z)}{U_N(x_t)} \leq 1 + \frac{2}{N e_K(t)} s_{K/N} \left(\frac{1}{2} d(x_0, z) \right)^2 \quad (2.17)$$

- (iii) *If U is bounded below we have the uniform continuity estimate*

$$s_{K/N} \left(\frac{1}{2} d(x_{t_1}, x_{t_0}) \right)^2 \leq \frac{N}{2e_{-K}(t_1 - t_0)} \left[1 - \frac{U_N(x_{t_0})}{U_N^{\max}} \right]. \quad (2.18)$$

Proof. By Lemma 2.8 (x_t) is an EVI_K flow of U and hence a metric gradient flow by [1, Proposition 3.9]. (2.17) follows immediately from (2.19) in Proposition 2.11 below by taking $t_0 = 0$. The uniform continuity estimate (2.18) is obtained similarly by taking $z = x_{t_0}$. □

The following result collects several equivalent reformulations of the definition of $\text{EVI}_{K,N}$ gradient flows which will be useful in the sequel. For this we say that a subset $D \subset D(U)$ is *dense in energy*, if for any $z \in D(U)$ there exists a sequence $(z_n) \subset D$ such that $d(z_n, z) \rightarrow 0$ and $U(z_n) \rightarrow U(z)$ as $n \rightarrow \infty$. Moreover, for a function $f : I \rightarrow \mathbb{R}$ on some interval I we use the notation

$$\frac{d^+}{dt}f(t) = \limsup_{h \searrow 0} \frac{f(t+h) - f(t)}{h}$$

to denote the right derivative.

Proposition 2.11. *Let $D \subset D(U)$ be dense in energy and let $x : (0, \infty) \rightarrow D(U)$ be a locally absolutely continuous curve with $\lim_{t \rightarrow 0} x_t = x_0$. Then (x_t) is an $\text{EVI}_{K,N}$ gradient flow of U if and only if one of the following statements holds:*

- (i) *The differential inequality (2.12) holds for all $z \in D$ and a.e. $t > 0$.*
- (ii) *For all $z \in D$ and all $0 \leq t_0 \leq t_1$:*

$$e_K(t_1 - t_0) \frac{N}{2} \left(1 - \frac{U_N(z)}{U_N(x_{t_1})} \right) \geq e^{K(t_1 - t_0)} s_{K/N} \left(\frac{1}{2} d(x_{t_1}, z) \right)^2 - s_{K/N} \left(\frac{1}{2} d(x_{t_0}, z) \right)^2. \quad (2.19)$$

- (iii) *For all $z \in D$ and all $t > 0$:*

$$\frac{d^+}{dt} s_{K/N} \left(\frac{1}{2} d(x_t, z) \right)^2 + K \cdot s_{K/N} \left(\frac{1}{2} d(x_t, z) \right)^2 \leq \frac{N}{2} \left(1 - \frac{U_N(z)}{U_N(x_t)} \right) \quad (2.20)$$

Proof. We prove the equivalence of Definition 2.7 and (ii). Assume that (x_t) is an $\text{EVI}_{K,N}$ flow and note that the right hand side of (2.12) can be rewritten as

$$e^{-Kt} \frac{d}{dt} \left[e^{Kt} s_{K/N} \left(\frac{1}{2} d(x_t, z) \right)^2 \right].$$

Integrating from t_0 to t_1 and using that the map $t \mapsto U_N(x_t)$ is non-decreasing by (i) of Proposition 2.10 then yields (2.19) for all $z \in D(U)$. Conversely, differentiating (2.19) yields (2.12). The fact that (2.19) holds for all $z \in D(U)$ if and only if it holds for all $z \in D$ is obvious.

Similar arguments show the equivalence of Definition 2.7 with (i) and (iii). \square

An important property of $\text{EVI}_{K,N}$ flows is the following contraction estimate.

Theorem 2.12. *Let $(x_t), (y_t)$ be two $\text{EVI}_{K,N}$ gradient flows of U starting from x_0 resp. y_0 . Then for all $s, t \geq 0$:*

$$s_{K/N} \left(\frac{1}{2} d(x_t, y_s) \right)^2 \leq e^{-K(s+t)} s_{K/N} \left(\frac{1}{2} d(x_0, y_0) \right)^2 + \frac{N}{K} \left(1 - e^{-K(s+t)} \right) \frac{(\sqrt{t} - \sqrt{s})^2}{2(s+t)}. \quad (2.21)$$

Proof. Let us fix $s, t > 0$. Choose $\lambda, r > 0$ such that $\lambda r = t$ and $\lambda^{-1} r = s$, i.e. $\lambda = \sqrt{\frac{t}{s}}$ and $r = \sqrt{ts}$. From (2.19) applied to (x_t) with $z = y_{\lambda^{-1}r}$ and $t_0 = \lambda r, t_1 = \lambda(r + \varepsilon)$ for some $\varepsilon > 0$ we obtain

$$\begin{aligned} \frac{N}{2} \frac{U_N(y_{\lambda^{-1}r})}{U_N(x_{\lambda(r+\varepsilon)})} &\leq \frac{N}{2} - \frac{1}{e^{-K(\lambda\varepsilon)}} s_{K/N} \left(\frac{1}{2} d(x_{\lambda(r+\varepsilon)}, y_{\lambda^{-1}r}) \right)^2 \\ &\quad + \frac{1}{e_K(\lambda\varepsilon)} s_{K/N} \left(\frac{1}{2} d(x_{\lambda r}, y_{\lambda^{-1}r}) \right)^2. \end{aligned} \quad (2.22)$$

Similarly, choosing $z = x_{\lambda(r+\varepsilon)}$ and $t_0 = \lambda^{-1}r, t_1 = \lambda^{-1}(r + \varepsilon)$ and applying (2.19) to (y_s) we obtain

$$\begin{aligned} \frac{N}{2} \frac{U_N(x_{\lambda(r+\varepsilon)})}{U_N(y_{\lambda^{-1}(r+\varepsilon)})} &\leq \frac{N}{2} - \frac{1}{e_{-K}(\lambda^{-1}\varepsilon)} s_{K/N} \left(\frac{1}{2} d(y_{\lambda^{-1}(r+\varepsilon)}, x_{\lambda(r+\varepsilon)}) \right)^2 \\ &\quad + \frac{1}{e_K(\lambda^{-1}\varepsilon)} s_{K/N} \left(\frac{1}{2} d(y_{\lambda^{-1}r}, x_{\lambda(r+\varepsilon)}) \right)^2. \end{aligned} \quad (2.23)$$

Multiplying (2.22) and (2.23) after taking square roots and using Young's inequality, $2\sqrt{ab} \leq \lambda a + \lambda^{-1}b$, we deduce the estimate

$$\begin{aligned} N \sqrt{\frac{U_N(y_{\lambda^{-1}r})}{U_N(y_{\lambda^{-1}(r+\varepsilon)})}} &\leq \frac{N}{2}(\lambda^{-1} + \lambda) \\ &\quad + s_{K/N} \left(\frac{1}{2} d(y_{\lambda^{-1}r}, x_{\lambda(r+\varepsilon)}) \right)^2 \left[\frac{\lambda^{-1}}{e_K(\lambda^{-1}\varepsilon)} - \frac{\lambda}{e_{-K}(\lambda\varepsilon)} \right] \\ &\quad + s_{K/N} \left(\frac{1}{2} d(x_{\lambda r}, y_{\lambda^{-1}r}) \right)^2 \left[\frac{\lambda}{e_K(\lambda\varepsilon)} - \frac{\lambda^{-1}}{e_{-K}(\lambda^{-1}\varepsilon)} \right] \\ &\quad - \frac{\lambda^{-1}\varepsilon}{e_{-K}(\lambda^{-1}\varepsilon)} \frac{1}{\varepsilon} \left[s_{K/N} \left(\frac{1}{2} d(y_{\lambda^{-1}(r+\varepsilon)}, x_{\lambda(r+\varepsilon)}) \right)^2 - s_{K/N} \left(\frac{1}{2} d(x_{\lambda r}, y_{\lambda^{-1}r}) \right)^2 \right]. \end{aligned} \quad (2.24)$$

Note that as $\varepsilon \rightarrow 0$ we have

$$\frac{e_{-K}(\lambda^{-1}\varepsilon)}{\lambda^{-1}\varepsilon} \rightarrow 1 \quad \text{and} \quad \left[\frac{\lambda^{-1}}{e_K(\lambda^{-1}\varepsilon)} - \frac{\lambda}{e_{-K}(\lambda\varepsilon)} \right] \rightarrow -\frac{K}{2}(\lambda + \lambda^{-1}).$$

Hence, if we consider the function $g : \mathbb{R}_+ \rightarrow \mathbb{R}$ given by

$$g(\tau) = \frac{2}{N} s_{K/N} \left(\frac{1}{2} d(x_{\lambda\tau}, y_{\lambda^{-1}\tau}) \right)^2$$

and take the limit as $\varepsilon \searrow 0$ in (2.24) we obtain

$$\frac{d^+}{d\tau} \Big|_{\tau=r} g(\tau) \leq -K(\lambda + \lambda^{-1})g(r) + (\lambda + \lambda^{-1} - 2).$$

By an application of Gronwall's lemma we deduce that

$$g(r) \leq e^{-K(\lambda + \lambda^{-1})r} \left[g(0) + \frac{\lambda + \lambda^{-1} - 2}{(\lambda + \lambda^{-1})} e_K((\lambda + \lambda^{-1})r) \right].$$

Rewriting r, λ in terms of s, t finally yields (2.21). \square

Remark 2.13. In the limit $d(x_0, y_0) \rightarrow 0$ and $s \rightarrow t$ the contraction estimate (2.21) reads asymptotically as follows:

$$\begin{aligned} d(x_t, y_s)^2 &\leq e^{-2Kt} d(x_0, y_0)^2 + \frac{N}{K} \frac{1 - e^{-2Kt}}{4t^2} \cdot |s - t|^2 \\ &\quad + o(d(x_0, y_0)^2 + |t - s|^2). \end{aligned} \quad (2.25)$$

Corollary 2.14. *For each $x_0 \in \overline{D(U)}$ there exist at most one $\text{EVI}_{K,N}$ gradient flow of U starting from x_0 . The maps $P_t : x_0 \mapsto x_t$, where (x_t) is the unique gradient flow starting from x_0 constitute a continuous semigroup defined on a closed (possibly empty) subset of $\overline{D(U)}$.*

2.3. $\text{EVI}_{K,N}$ implies (K, N) -convexity. We now investigate the relation between the Evolution Variational Inequality and geodesic convexity of the functional U .

Definition 2.15. *Let $K, N \in \mathbb{R}$ with $N > 0$. We say that the functional U is (K, N) -convex if and only if for each pair $x_0, x_1 \in D(U)$ there exists a constant speed geodesic $\gamma : [0, 1] \rightarrow X$ connecting x_0 to x_1 such that for all $t \in [0, 1]$:*

$$U_N(\gamma_t) \geq \sigma_{K/N}^{(1-t)}(d) \cdot U_N(\gamma_0) + \sigma_{K/N}^{(t)}(d) \cdot U_N(\gamma_1), \quad (2.26)$$

where $d = d(\gamma_0, \gamma_1)$.

If (2.26) holds for every geodesic $\gamma : [0, 1] \rightarrow D(U)$ we say that U is strongly (K, N) -convex.

Theorem 2.16. *Assume that for every starting point $x_0 \in \overline{D(U)}$ the $\text{EVI}_{K,N}$ flow for U exists. Then U is strongly (K, N) -convex.*

Proof. Let P denote the $\text{EVI}_{K,N}$ gradient flow semigroup of U . We treat the case $K \neq 0$ first. So let $(\gamma_s)_{s \in [0,1]}$ be a constant speed geodesic. Let us fix $s \in [0, 1], t > 0$ and set $\gamma_s^t := P_t \gamma_s$. We can assume that $d := d(\gamma_0, \gamma_1) \neq 0$. Using the identity (2.14) we see that (2.19) can be rewritten as

$$\begin{aligned} \frac{U_N(z)}{U_N(P_{t_1}x)} e_K(t_1 - t_0) &\leq \frac{1}{K} \left[e^{K(t_1 - t_0)} c_{K/N}(d(P_{t_1}x, z)) \right. \\ &\quad \left. - c_{K/N}(d(P_{t_0}x, z)) \right]. \end{aligned} \quad (2.27)$$

Using (2.27) with $t_0 = 0, t_1 = t, x = \gamma_s$ and $z = \gamma_0$ respectively $z = \gamma_1$ we immediately obtain

$$\begin{aligned} &\sigma_{K/N}^{(1-s)}(d) \cdot U_N(\gamma_0) + \sigma_{K/N}^{(s)}(d) \cdot U_N(\gamma_1) \\ &\leq \frac{U_N(\gamma_s^t)}{K \cdot e_K(t)} \left[\sigma_{K/N}^{(1-s)}(d) \cdot \left(e^{Kt} c_{K/N}(d(\gamma_s^t, \gamma_0)) - c_{K/N}(d(\gamma_s, \gamma_0)) \right) \right. \\ &\quad \left. + \sigma_{K/N}^{(s)}(d) \cdot \left(e^{Kt} c_{K/N}(d(\gamma_s^t, \gamma_1)) - c_{K/N}(d(\gamma_s, \gamma_1)) \right) \right]. \end{aligned}$$

Let A denote the term in square brackets in the last inequality. The claim follows if we show that for t small enough we have $A \leq K \cdot e_K(t) = e^{Kt} - 1$ if $K > 0$ and $A \geq e^{Kt} - 1$ if $K < 0$. Using the fact that $d(\gamma_s, \gamma_{s'}) = |s - s'|d$, we first find

$$\begin{aligned} A &= \frac{e^{Kt}}{s_{K/N}(d)} \left[s_{K/N}((1-s)d) \cdot c_{K/N}(d(\gamma_s^t, \gamma_0)) + s_{K/N}(sd) \cdot c_{K/N}(d(\gamma_s^t, \gamma_1)) \right] \\ &\quad - \frac{1}{s_{K/N}(d)} \left[s_{K/N}((1-s)d) \cdot c_{K/N}(sd) + s_{K/N}(sd) \cdot c_{K/N}((1-s)d) \right] \\ &:= A_1 + A_2. \end{aligned}$$

By the angle sum identity for \sin (resp. \sinh) we have $A_2 = -1$. To see that $A_1 \leq e^{Kt}$ (resp. $A_1 \geq e^{Kt}$), we observe the following fact, which is easily verified using the angle sum identities for trigonometric or hyperbolic functions: If $\alpha, \alpha' \geq 0$ and $\varepsilon, \varepsilon' \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ such that $\varepsilon + \varepsilon' \geq 0$, then, putting $\beta = \alpha + \varepsilon, \beta' = \alpha' + \varepsilon'$, we have that

$$\begin{aligned} \sin(\alpha) \cos(\beta') + \cos(\beta) \sin(\alpha') &\leq \sin(\alpha + \alpha'), \\ \sinh(\alpha) \cosh(\beta') + \cosh(\beta) \sinh(\alpha') &\geq \sinh(\alpha + \alpha'). \end{aligned}$$

To conclude, we apply this with $\alpha = (1-s)d, \alpha' = sd$ and $\varepsilon = d(\gamma_s^t, \gamma_1) - (1-s)d, \varepsilon' = d(\gamma_s^t, \gamma_0) - sd$ and note that $\varepsilon + \varepsilon' \geq 0$ by the triangle inequality.

Finally, we treat the case $K = 0$. By Lemma 2.8 P is a $\text{EVI}_{K',N}$ flow for every $K' < 0$. Thus by the previous argument (2.26) holds with K' instead of K and we can pass to the limit as $K' \nearrow 0$. \square

We conclude this section with some remarks about (K, N) -convexity. The first property is immediate from the definition.

Lemma 2.17. *If $U : X \rightarrow (-\infty, \infty]$ is (K, N) -convex, then for $\lambda > 0$ the functional $\lambda \cdot U$ is $(\lambda K, \lambda N)$ -convex.*

Lemma 2.18. *Let $U^i : X \rightarrow (-\infty, \infty]$ be strongly (K_i, N_i) -convex functionals for $i = 1, 2$ and real numbers $K_i, N_i \in \mathbb{R}$ with $N_i > 0$. Then the functional $U := U^1 + U^2$ is strongly $(K_1 + K_2, N_1 + N_2)$ -convex.*

Proof. Let us set $K = K_1 + K_2$ and $N = N_1 + N_2$ and fix a constant speed geodesic $\gamma : [0, 1] \rightarrow D(U) = D(U^1) \cap D(U^2)$. By the convexity assumption on U^1 and U^2 we have

$$\begin{aligned} \log U_N(\gamma_t) &= \frac{N_1}{N} \frac{(-1)}{N_1} U^1(\gamma_t) + \frac{N_2}{N} \frac{(-1)}{N_2} U^2(\gamma_t) \\ &\geq \frac{N_1}{N} G\left(\frac{(-1)}{N_1} U^1(\gamma_0), \frac{(-1)}{N_1} U^1(\gamma_1), \frac{K_1}{N_1}\right) \\ &\quad + \frac{N_2}{N} G\left(\frac{(-1)}{N_2} U^2(\gamma_0), \frac{(-1)}{N_2} U^2(\gamma_1), \frac{K_2}{N_2}\right), \end{aligned}$$

where the function G is given by (2.28) with $\theta = d(\gamma_0, \gamma_1)$. By Lemma 2.19 below, G is convex. Hence we obtain

$$\log U_N(\gamma_t) \geq G\left(\frac{(-1)}{N} U(\gamma_0), \frac{(-1)}{N} U(\gamma_1), \frac{K}{N}\right).$$

Taking the exponential on both sides yields the claim. \square

Lemma 2.19. *For any fixed $\theta \geq 0$ and $t \in [0, 1]$ the function $G : \mathbb{R} \times \mathbb{R} \times (-\infty, \pi^2) \rightarrow \mathbb{R}$ given by*

$$G(x, y, \kappa) = \log \left[\sigma_{\kappa}^{(1-t)}(\theta) e^x + \sigma_{\kappa}^{(t)}(\theta) e^y \right] \quad (2.28)$$

is convex.

Proof. We define the function $g^{(t)} : \kappa \mapsto \log \sigma_{\kappa}^{(t)}(\theta)$ on $(-\infty, \pi^2)$ and write

$$G(x, y, z) = F\left(x + g^{(1-t)}(\kappa), y + g^{(t)}(\kappa)\right),$$

where $F(u, v) = \log(e^u + e^v)$. The claim then follows by noting that the function F is convex and increasing on $\mathbb{R} \times \mathbb{R}$ and that the functions $g^{(t)}$ are convex. \square

Finally we remark that the notion of (K, N) -convexity is consistent in the parameters K and N .

Lemma 2.20. *If U is (K, N) -convex for some numbers $K, N \in \mathbb{R}$ with $N > 0$ then it is also (K', N') -convex for all $K' \leq K$ and $N' \geq N$. Moreover, it is K -convex in the sense that for each pair $x_0, x_1 \in D(U)$ there exist a constant speed geodesic $\gamma : [0, 1] \rightarrow X$ connecting x_0 to x_1 such that for all $t \in [0, 1]$:*

$$U(\gamma_t) \leq (1-t)U(\gamma_0) + tU(\gamma_1) - \frac{K}{2}t(1-t)d(\gamma_0, \gamma_1)^2. \quad (2.29)$$

Proof. Consistency in K is immediate from the fact that for any fixed t and θ the coefficient $\sigma_{K/N}^{(t)}(\theta)$ is increasing in K . Consistency in N is a consequence e.g. of Lemma 2.18 and the trivial observation that for any $N' > N$ the constant functional $U^0 \equiv 0$ is $(0, N' - N)$ -convex.

Using the consistency in N we can derive (2.29) by subtracting 1 on both sides of (2.26), multiplying with N and passing to the limit $N \nearrow \infty$. Here we use the fact that $\sigma_{K/N}^{(t)}(\theta) = t + o(1/N)$ and $U_N(x) = 1 - U(x)/N + o(1/N)$. \square

3. ENTROPIC AND RIEMANNIAN CURVATURE-DIMENSION CONDITIONS

3.1. The entropic curvature-dimension condition. In this section we introduce a new curvature-dimension condition for metric measure spaces based on (K, N) -convexity of the entropy on the Wasserstein space.

Let (X, d, m) be a metric measure space, i.e. (X, d) is a complete and separable metric space and m is a locally finite Borel measure on X . We denote by $\mathcal{P}_2(M, d)$ the L^2 -Wasserstein space over (X, d) , i.e. the set of all Borel probability measures μ satisfying

$$\int d(x_0, x)^2 \mu(dx) < \infty$$

for some, hence any, $x_0 \in X$. The subspace of all measures absolutely continuous w.r.t. m is denoted by $\mathcal{P}_2(X, d, m)$. The L^2 -Wasserstein distance between $\mu_0, \mu_1 \in \mathcal{P}_2(X, d)$ is defined by

$$W_2(\mu_0, \mu_1)^2 = \inf \int d(x, y)^2 dq(x, y),$$

where the infimum is taken over all Borel probability measures q on $X \times X$ with marginals μ_0 and μ_1 .

Given a measure $\mu \in \mathcal{P}_2(X, d)$ we define its relative entropy by

$$\text{Ent}(\mu) := \int \rho \log \rho \, dm,$$

if $\mu = \rho m$ is absolutely continuous w.r.t. m and $(\rho \log \rho)_+$ is integrable. Otherwise we set $U(\mu|m) = +\infty$. The subset of probability measures with finite entropy will be denoted by $\mathcal{P}_2^*(X, d, m)$. Moreover, for a number $N \in \mathbb{R}$ with $N > 0$ we introduce the functional $U_N : \mathcal{P}_2(X, d) \rightarrow [0, \infty]$ by

$$U_N(\mu) := \exp\left(-\frac{1}{N} \text{Ent}(\mu)\right).$$

Definition 3.1. *Given two numbers $K, N \in \mathbb{R}$ with $N > 0$ we say that a metric measure space (X, d, m) satisfies the entropic curvature-dimension condition $\text{CD}^e(K, N)$ if and only if for each pair $\mu_0, \mu_1 \in \mathcal{P}_2^*(X, d, m)$ there exists a constant speed geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_2^*(X, d, m)$ connecting μ_0 to μ_1 such that for all $t \in [0, 1]$:*

$$U_N(\Gamma(t)) \geq \sigma_{K/N}^{(1-t)}(W_2(\mu_0, \mu_1))U_N(\mu_0) + \sigma_{K/N}^{(t)}(W_2(\mu_0, \mu_1))U_N(\mu_1). \quad (3.1)$$

If (3.1) holds for any constant speed geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_2^*(X, d, m)$ we say that (X, d, m) is a strong $\text{CD}^e(K, N)$ space.

Remark 3.2. In other words, the $\text{CD}^e(K, N)$ -condition means that the entropy is (K, N) -convex along Wasserstein geodesic. In Section 2.2 the notion of (K, N) -convexity was defined for lower semi-continuous functions with values in $(-\infty, \infty]$. Note that the volume growth estimate in $\text{CD}^e(K, N)$ spaces given by Proposition 3.7 below implies that for a suitable constant $c > 0$ and any $x_0 \in X$

$$\int \exp(-cd(x_0, x)^2 m(dx)) < \infty.$$

It is well known that the latter implies that Ent does not take the value $-\infty$ on $\mathcal{P}_2(X, d)$ and is lower semi-continuous w.r.t. W_2 . Thus Definition 3.1 fits well into the setting of Section 2.2.

As an immediate consequence of Lemma 2.20 we obtain the following consistency result.

Lemma 3.3. *Assume that (X, d, M) satisfies the $\text{CD}^e(K, N)$ condition for numbers $K, N \in \mathbb{R}$ with $N > 0$. Then it also satisfies $\text{CD}^e(K', N')$ for any $K' \leq K$ and $N' \geq N$. Moreover, it satisfies the $\text{CD}(K, \infty)$ condition.*

Using similar arguments as in the case of the $\text{CD}(K, \infty)$ condition introduced in [24] it is immediate to check that $\text{CD}^e(K, N)$ is invariant under isomorphisms of metric measure spaces. Moreover, adapting [24, Thm. 4.20], one can check that it is stable under convergence of metric measure spaces in the transportation distance \mathbb{D} , also introduced in [24].

As an application of the additivity of (K, N) -convexity we note the following

Proposition 3.4 (Weighted spaces). *Let (X, d, m) be a geodesic metric measure space satisfying $\text{CD}^e(K, N)$ and let $V : X \rightarrow \mathbb{R}$ be a measurable function bounded from below that is (K', N') -convex in the sense of Definition 2.15. Then $(X, d, e^{-V}m)$ satisfies $\text{CD}^e(K + K', N + N')$.*

Proof. Using [24, Lemma 2.11], it is immediate to check that the functional $\bar{V} : \mathcal{P}_2(X, d) \rightarrow (-\infty, \infty]$ defined by $\bar{V}(\mu) = \int V \, d\mu$ is (K', N') -convex on $\mathcal{P}_2(X, d)$. By the lower boundedness of V we have $\mathcal{P}_2(X, d, e^{-V}m) \subset \mathcal{P}_2(X, d, m)$. Now the assertion is a consequence of the observation that

$$\text{Ent}(\mu|e^{-V}m) = \text{Ent}(\mu|m) + \bar{V}(\mu).$$

and Lemma 2.18. \square

We will now derive some first geometric consequences of the entropic curvature-dimension condition.

Proposition 3.5 (Generalized Brunn–Minkowski inequality). *Assume that (X, d, m) satisfies the condition $\text{CD}^e(K, N)$ for real numbers K, N with $N \geq 1$. Then for all measurable sets $A_0, A_1 \subset X$ with $m(A_0), m(A_1) > 0$ and all $t \in [0, 1]$ we have*

$$m(A_t)^{1/N} \geq \sigma_{K/N}^{(1-t)}(\Theta) \cdot m(A_0)^{1/N} + \sigma_{K/N}^{(t)}(\Theta) \cdot m(A_1)^{1/N}, \quad (3.2)$$

where A_t denotes the set t -midpoints and Θ the minimal/maximal distance between points in A_0 and A_1 , i.e.

$$\begin{aligned} A_t &= \{ \gamma_t : \gamma : [0, 1] \rightarrow X \text{ geodesic s.t. } \gamma_0 \in A_0, \gamma_1 \in A_1 \}, \\ \Theta &= \begin{cases} \inf_{x_0 \in A_0, x_1 \in A_1} d(x_0, x_1), & K \geq 0, \\ \sup_{x_0 \in A_0, x_1 \in A_1} d(x_0, x_1), & K < 0. \end{cases} \end{aligned}$$

Proof. We first prove the assertion under the assumption that $m(A_0), m(A_1) < \infty$, the general case then follows by approximating the sets A_0, A_1 by sets of finite volume. Applying the condition $\text{CD}^e(K, N)$ to $\mu_i = m(A_i)^{-1} \mathbf{1}_{A_i} m$ for $i = 0, 1$ yields

$$U_N(\Gamma_t) \geq \sigma_{K/N}^{(1-t)}(W_2(\mu_0, \mu_1)) \cdot m(A_0)^{1/N} + \sigma_{K/N}^{(t)}(W_2(\mu_0, \mu_1)) \cdot m(A_1)^{1/N}, \quad (3.3)$$

where $\Gamma_t = \rho_t m$ is the t -midpoint of a geodesic Γ connecting μ_0 and μ_1 . Since $\Gamma(t)$ is concentrated on A_t , a double application of Jensen's inequality gives that

$$\begin{aligned} U_N(\Gamma_t) &= \exp \left(-\frac{1}{N} \int \rho_t \, d\Gamma_t \right) \leq \int \rho_t^{-1/N} \, d\Gamma_t \\ &= \int_{A_t} \rho_t^{1-1/N} \, dm \leq m(A_t)^{1/N}. \end{aligned}$$

Hence (3.2) follows by noting that $\theta \mapsto \sigma_{K/N}^{(t)}(\theta)$ is increasing if $K \geq 0$ and decreasing if $K < 0$ and that $W_2(\mu_0, \mu_1) \geq \Theta$ (resp. $\leq \Theta$). \square

The Brunn–Minkowski inequality entails further geometric consequences like a Bishop–Gromov type volume growth estimate and a generalized Bonnet–Myers theorem. The following results can be deduced from Proposition 3.5 using similar arguments as in [24] and replacing the coefficients $\tau_{K/N}^{(t)}(\cdot)$ by $\sigma_{K/N}^{(t)}(\cdot)$.

Remark 3.6. The estimates presented below are not sharp, yet they provide necessary local compactness results for example. We will see below that under the assumption that (X, d, m) is non-branching the $\text{CD}^e(K, N)$ condition is equivalent to the $\text{CD}^*(K, N)$ condition. It has been proven by Cavaletti and Sturm [10] that under the same assumption $\text{CD}^*(K, N)$ implies the measure contraction property $\text{MCP}(K, N)$ from which a sharp Bishop–Gromov and Lichnerowicz inequality can be derived, see [24].

To state the volume growth estimate we introduce the following notation. Given a metric measure space (X, d, m) and a point $x_0 \in \text{supp}[m]$ we denote by

$$v(r) := m(\overline{B_r(x_0)})$$

the volume of the closed ball of radius r around x_0 . Moreover, we set

$$s(r) := \limsup_{\delta \rightarrow 0} \frac{1}{\delta} m(\overline{B_{r+\delta}(x_0)} \setminus B_r(x_0))$$

for the volume of the corresponding sphere.

Proposition 3.7 (Generalized Bishop–Gromov inequality). *Assume that (X, d, m) satisfies the condition $\text{CD}^e(K, N)$ with $K, N \in \mathbb{R}$ and $N \geq 1$. Then each bounded closed set $M \subset \text{supp}[m]$ is compact and has finite volume. More precisely, if $K > 0$, for each $x_0 \in \text{supp}[m]$ and $0 < r < R \leq \pi\sqrt{K/N}$,*

$$\frac{s(r)}{s(R)} \geq \left(\frac{\sin(r\sqrt{K/N})}{\sin(R\sqrt{K/N})} \right)^N \quad \text{and} \quad \frac{v(r)}{v(R)} \geq \frac{\int_0^r \sin(t\sqrt{K/N})^N dt}{\int_0^R \sin(t\sqrt{K/N})^N dt}. \quad (3.4)$$

In the case $K \leq 0$ analogous estimates hold with an appropriate interpretation of the right-hand side.

Corollary 3.8 (Generalized Bonnet–Myers theorem). *If (X, d, m) satisfies the condition $\text{CD}^e(K, N)$ with $K > 0$ and $N \geq 1$, then the support of m is compact and its diameter L can be bounded as $L \leq \pi\sqrt{K/N}$.*

It turns out that under mild assumptions the modified curvature-dimension condition $\text{CD}^e(K, N)$ is equivalent to the reduced curvature-dimension condition $\text{CD}^*(K, N)$ introduced in [8]. We recall here the definition. Denote by $\mathcal{P}_\infty(X, d, m)$ the set of measures in $\mathcal{P}_2(X, d, m)$ with bounded support.

Definition 3.9. *We say that a metric measure space (X, d, m) satisfies the reduced curvature-dimension condition $\text{CD}^*(K, N)$ if and only if for each pair $\mu_0 = \rho_0 m, \mu_1 = \rho_1 m \in \mathcal{P}_\infty(X, d, m)$ there exist an optimal coupling q of them and a geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_\infty(X, d, m)$ connecting them such that for all $t \in [0, 1]$ and $N' \geq N$:*

$$\int \rho_t^{-\frac{1}{N'}} d\Gamma(t) \geq \int_{X \times X} \left[\sigma_{K/N'}^{(1-t)}(d(x_0, x_1)) \rho_0(x_0)^{-\frac{1}{N'}} + \sigma_{K/N'}^{(t)}(d(x_0, x_1)) \rho_1(x_1)^{-\frac{1}{N'}} \right] dq(x_0, x_1). \quad (3.5)$$

If (3.5) holds for any geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_\infty(X, d, m)$ we say that (X, d, m) is a strong $\text{CD}^(K, N)$ space.*

We first note the following adaption of a result from [24].

Lemma 3.10. *Assume that (X, d, m) is non-branching and satisfies $\text{CD}^e(K, N)$ for some $K, N \in \mathbb{R}$ with $N > 0$. Then for every $x \in \text{supp}[m]$ and m -a.e. $y \in X$ there exists a unique geodesic connecting x to y .*

Moreover, there exists a measurable map $\gamma : X^2 \rightarrow \text{Geo}(X)$ such that for $m \otimes m$ -a.e. $(x, y) \in X^2$ the curve $t \mapsto \gamma_t(x, y)$ is the unique geodesic connecting x and y .

Proof. The proof follows from the very same arguments as [24, Lemma II4.1] where $\text{CD}(K, N)$ is assumed in stead of $\text{CD}^e(K, N)$. We just have to note that the curvature-dimension condition only enters through a Brunn–Minkowski inequality which holds in our setting according to Proposition 3.5. \square

Theorem 3.11. *Let (X, d, m) be a non-branching metric measure space. Then the following assertions are equivalent:*

- (i) (X, d, m) satisfies $\text{CD}^*(K, N)$,
- (ii) For each pair $\mu_0, \mu_1 \in \mathcal{P}_\infty(X, d, m)$ and each optimal coupling q of them we have

$$\rho_t(\gamma_t(x_0, x_1))^{-\frac{1}{N}} \geq \sigma_{K/N}^{(1-t)}(d(x_0, x_1)) \rho_0(x_0)^{-\frac{1}{N}} + \sigma_{K/N}^{(t)}(d(x_0, x_1)) \rho_1(x_1)^{-\frac{1}{N}}, \quad (3.6)$$

for q -a.e. $(x_0, x_1) \in X \times X$, where ρ_t denotes the density of the push forward of q under the map $(x_0, x_1) \mapsto \gamma_t(x_0, x_1)$,

- (iii) (X, d, m) satisfies $\text{CD}^e(K, N)$.

Proof. The equivalence of (i) and (ii) has already been proven in [8, Proposition 2.8].

(ii) \Rightarrow (iii): First note that by an approximation argument as in [8, Lemma 2.11] one can show that (3.6) also holds for $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ not necessarily with bounded support. Now fix $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ and an optimal coupling q of them. Taking logarithms on both sides of (3.6) we obtain

$$-\frac{1}{N} \log \rho_t(\gamma_t(x_0, x_1)) \geq G\left(-\frac{1}{N} \log \rho_0(x_0), -\frac{1}{N} \log \rho_1(x_1), d(x_0, x_1)^2\right), \quad (3.7)$$

where the function G is given by

$$G(x, y, z) := \log \left[\sigma_{K/N}^{(1-t)}(\sqrt{z}) e^x + \sigma_{K/N}^{(t)}(\sqrt{z}) e^y \right]. \quad (3.8)$$

From Lemma 2.19 and the fact that $\sigma_{K/N}^{(1-t)}(\sqrt{z}) = \sigma_{zK/N}^{(1-t)}(1)$ we deduce that G is convex. Integrating (3.7) w.r.t. q and using Jensen's inequality we obtain

$$-\frac{1}{N} \text{Ent}(\Gamma(t)) \geq G\left(-\frac{1}{N} \text{Ent}(\mu_0), -\frac{1}{N} \text{Ent}(\mu_1), W_2(\mu_0, \mu_1)^2\right),$$

where $\Gamma(t)$ denotes the measure $\rho_t m$. Hence (3.1) follows by taking the exponential on both sides.

(iii) \Rightarrow (ii): Here we follow closely the arguments in the proof of [24, Prop. 4.1]. Fix $\mu_0, \mu_1 \in \mathcal{P}_\infty(X, d, m)$ and an optimal coupling q of them. Let $\{M_n\}_{n \in \mathbb{N}}$ be a \cap -stable generator of the Borel σ -field of X with $m(\partial M_n) = 0$ for all n . For each n consider the disjoint covering of X given by the 2^n sets $L_1 = M_1 \cap \dots \cap M_n$, $L_2 = M_1 \cap \dots \cap M_n^c$, \dots , $L_{2^n} = M_1^c \cap \dots \cap M_n^c$. For fixed n and $i, j = 1, \dots, 2^n$ we define sets $A_{i,j} = L_i \times L_j$ and probability measures $\mu_0^{i,j}, \mu_1^{i,j}$ by

$$\mu_0^{i,j}(B) = \alpha_{i,j}^{-1} q((B \cap L_i) \times L_j), \quad \mu_1^{i,j}(B) = \alpha_{i,j}^{-1} q(L_i \times (B \cap L_j)),$$

provided that $\alpha_{i,j} = q(A_{i,j}) > 0$. By (iii) we can choose optimal couplings $q^{i,j}$ of them such that

$$\begin{aligned} U_N(\mu_t^{i,j}) &\geq \sigma_{K/N}^{(1-t)}(W_2(\mu_0^{i,j}, \mu_1^{i,j})) \cdot U_N(\mu_0^{i,j}) \\ &\quad + \sigma_{K/N}^{(t)}(W_2(\mu_0^{i,j}, \mu_1^{i,j})) \cdot U_N(\mu_1^{i,j}), \end{aligned} \quad (3.9)$$

where $\mu_t^{i,j} = (\gamma_t)_\# q^{i,j}$ and $\gamma : X^2 \rightarrow \text{Geo}(X)$ is the map from Lemma 3.10. Define a coupling of the measures μ_0, μ_1 by

$$q^{(n)} := \sum_{i,j=1}^{2^n} \alpha_{i,j} q^{i,j}.$$

Since the measures $q^{i,j}$ are mutually singular and X is non-branching, also the measures $\mu_t^{i,j}$ are mutually singular. Setting $\mu_t^{(n)} = (\gamma_t)_\# q^{(n)}$, we conclude that $\rho_t^{(n)} \circ \gamma_t = \alpha^{i,j} \rho_t^{i,j} \circ \gamma_t$ on the set $A_{i,j}$. Plugging this into (3.9) and taking logarithms on both sides we find

$$\begin{aligned} &-\frac{\alpha_{i,j}^{-1}}{N} \int_{A_{i,j}} \log \rho_t^{(n)}(\gamma_t) dq^{(n)} \\ &\geq G\left(-\frac{\alpha_{i,j}^{-1}}{N} \int_{A_{i,j}} \log \rho_0 dq^{(n)}, -\frac{\alpha_{i,j}^{-1}}{N} \int_{A_{i,j}} \log \rho_1 dq^{(n)}, \alpha_{i,j}^{-1} \int_{A_{i,j}} d^2 dq^{(n)}\right), \end{aligned} \quad (3.10)$$

where G is defined in (3.8). Since μ_0, μ_1 have bounded support, all the measures $q^{(n)}$ are supported in $B \times B$ for a suitable closed bounded set $B \subset X$. By Proposition 3.7 B is compact and has finite mass. Hence the couplings $q^{(n)}$ converge weakly, up to extraction of a subsequence, to an optimal coupling \tilde{q} of μ_0 and μ_1 . Since $m(\partial M_i) = 0$ for all i we deduce that

$$q(M_i \times M_j) = \lim_{n \rightarrow \infty} q^n(M_i \times M_j) = \tilde{q}(M_i \times M_j)$$

for all $i, j \in \mathbb{N}$ and hence $\tilde{q} = q$. By weak lower semi-continuity of the entropy we can pass to the limit as $n \rightarrow \infty$ in the left hand side of (3.10). Invoking furthermore the convexity of G given by Lemma 2.19 and Jensen's inequality we see that

$$\begin{aligned} & -\frac{\alpha^{-1}}{N} \int_A \log \rho_t(\gamma_t) \, dq \\ & \geq G\left(-\frac{\alpha^{-1}}{N} \int_A \log \rho_0 \, dq, -\frac{\alpha^{-1}}{N} \int_A \log \rho_1 \, dq, \alpha^{-1} \int_A d^2 \, dq\right), \end{aligned} \quad (3.11)$$

for any set A which is a union of a finite number of the sets $A_{i,j}$ and $\alpha = q(A)$. This implies the q -a.s. inequality (3.6). \square

Corollary 3.12. *For a metric measure space (X, d, m) the following assertions are equivalent:*

- (i) (X, d, m) is a strong $\text{CD}^*(K, N)$ space,
- (ii) For each pair $\mu_0, \mu_1 \in \mathcal{P}_\infty(X, d, m)$ (3.6) holds,
- (iii) (X, d, m) is a strong $\text{CD}^e(K, N)$ space.

Proof. Note that both (i) and (iii) imply that (X, d, m) satisfies the strong $\text{CD}(K, \infty)$ condition. Rajala and Sturm in [23] proved that in a strong $\text{CD}(K, \infty)$ space any dynamic optimal coupling of two absolutely continuous measures $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$ is supported on a set of non-branching geodesics. Thus the assertion follows from the same arguments as in the proof of Theorem 3.11. \square

3.2. Calculus and heat flow on metric measure spaces. Here we recapitulate briefly some of the results obtained by Ambrosio, Gigli and Savaré in a series of recent works, see [4, 5, 6]. In particular, we introduce notation and concepts that we use in the sequel about the powerful machinery of calculus on metric measure spaces developed by these authors. We refer to [4, 5] for more details on the definitions and results.

Let $(X, d, m) \in \mathcal{X}$ be a metric measure space. The basic object of study, introduced in [4] is the Cheeger energy. For a measurable function $f : X \rightarrow \mathbb{R}$ it can be defined by

$$\text{Ch}(f) = \frac{1}{2} \int |\nabla f|_w^2 \, dm,$$

where $|\nabla f|_w : X \rightarrow [0, \infty]$ denotes the so called minimal weak upper gradient of f . An important approximation result [4, Thm. 6.2] states that for $f \in L^2(X, m)$ the Cheeger energy can also be obtained by a relaxation procedure:

$$\text{Ch}(f) = \inf \left\{ \liminf_{n \rightarrow \infty} \frac{1}{2} \int |\nabla f_n|^2 \, dm \right\},$$

where the infimum is taken over all sequences of Lipschitz functions (f_n) converging to f in $L^2(X, m)$ and where $|\nabla f_n|$ denotes the local Lipschitz constant. In particular, Lipschitz functions are dense in the domain of Ch in $L^2(X, m)$ denoted by $D(\text{Ch}) = W^{1,2}(X, d, m)$.

It turns out that Ch is a convex and lower semi-continuous functional on $L^2(X, m)$. It allows to define the Laplacian $-\Delta f \in L^2(X, m)$ of a function $f \in W^{1,2}(X, d, m)$ as the element of minimal L^2 -norm in the subdifferential $\partial^- \text{Ch}(f)$ provided the latter is non-empty. In this generality, Ch is not necessarily a quadratic form and consequently Δ need not be a linear operator.

The classical theory of gradient flows of convex functionals in Hilbert-spaces allows to study the gradient flow of Ch in $L^2(X, m)$: For any $f \in L^2(X, m)$ there exists a unique continuous curve $(f_t)_{t \in [0, \infty)}$ in $L^2(X, m)$, locally absolutely continuous in $(0, \infty)$ with $f_0 = f$ such that $\frac{d}{dt} f_t \in \partial^- \text{Ch}(f_t)$ for a.e. $t > 0$. In fact, we have $f_t \in D(\Delta)$ and

$$\frac{d^+}{dt} f_t = \Delta f_t$$

for all $t > 0$. This gives rise to a semigroup $(H_t)_{t \geq 0}$ on $L^2(X, m)$ defined by $H_t f = f_t$, where f_t is the unique L^2 -gradient flow of Ch .

On the other hand, one can study the metric gradient flow of the relative entropy Ent in $\mathcal{P}_2(X, d)$. Under the assumption that (X, d, m) satisfies $\text{CD}(K, \infty)$ it has been proven in [12] and more generally in [4, Thm. 9.3(ii)] that for any $\mu \in D(\text{Ent})$ there exist a unique gradient flow of Ent starting from μ in the sense of Definition 2.6. This gives rise to a semigroup $(\mathcal{H}_t)_{t \geq 0}$ on $\mathcal{P}_2(X, d)$ defined by $\mathcal{H}_t \mu = \mu_t$ where μ_t is the unique gradient flow of Ent starting from μ .

One of the main result of [4] is the identification of the two gradient flows, which allows to consistently define the heat flow on $\text{CD}(K, \infty)$ spaces.

Theorem 3.13 ([4, Thm. 9.3]). *Let (X, d, m) be a $\text{CD}(K, \infty)$ space and let $f \in L^2(X, d, m)$ such that $\mu = fm \in \mathcal{P}_2(X, d)$. Then we have*

$$\mathcal{H}_t \mu = (\text{H}_t f) m \quad \forall t \geq 0.$$

A byproduct of this result is a representation of the slope of the entropy.

$$|\nabla^- \text{Ent}|(\rho m) = 4 \int |\nabla \sqrt{\rho}|_w^2 dm \quad (3.12)$$

for all probability densities ρ with $\sqrt{\rho} \in D(\text{Ch})$.

A basic property of the heat flow is the maximum principle, see [4, Thm. 4.16]: If $f \in L^2(X, m)$ satisfies $f \leq C$ m -a.e. then also $\text{H}_t f \leq C$ m -a.e. for all $t \geq 0$.

If Ch is assumed to be a quadratic form, and without any curvature assumption, the notion of weak upper gradient gives rise to a powerful calculus, in which not only the norm of the gradient, but also scalar products between gradients are defined, we refer to [5, Sec. 4.3] for details. We note briefly that given $f, g \in D(\text{Ch})$, the limit

$$\langle \nabla f, \nabla g \rangle := \lim_{\varepsilon \searrow 0} \frac{1}{2\varepsilon} (|\nabla f + \varepsilon g|_w^2 - |\nabla f|_w^2) \quad (3.13)$$

can be shown to exist in $L^1(X, m)$. Moreover, the map $D(\text{Ch})^2 \ni (f, g) \mapsto \langle \nabla f, \nabla g \rangle \in L^1(X, m)$ is bilinear, symmetric and satisfies

$$|\langle \nabla f, \nabla g \rangle| \leq |\nabla f|_w |\nabla g|_w.$$

For all $f, g, h \in D(\text{Ch}) \cap L^\infty(X, m)$ we have the Leibnitz rule:

$$\int \langle \nabla f, \nabla(gh) \rangle dm = \int h \langle \nabla f, \nabla g \rangle dm + \int g \langle \nabla f, \nabla h \rangle dm. \quad (3.14)$$

A quadratic Cheeger energy gives rise to a strongly local Dirichlet form $(\mathcal{E}, D(\mathcal{E}))$ on $L^2(X, m)$ by setting $\mathcal{E}(f, f) = \text{Ch}(f)$ and $D(\mathcal{E}) = W^{1,2}(X, d, m)$. In this case H_t is a semigroup of self-adjointed linear operators on $L^2(X, m)$ with the Laplacian Δ as its generator. The previous result implies that for $f, g \in W^{1,2}(X, d, m)$

$$\mathcal{E}(f, g) = \int \langle \nabla f, \nabla g \rangle dm,$$

i.e. the energy measure of \mathcal{E} has a density given by (3.13). Moreover, for $f \in W^{1,2}$ and $g \in D(\Delta)$ we have the integration by parts formula

$$\int \langle \nabla f, \nabla g \rangle dm = - \int f \Delta g dm. \quad (3.15)$$

3.3. The Riemannian curvature-dimension condition. In this section we introduce the notion of Riemannian curvature-dimension bounds. This notion can be seen as a generalization of the Riemannian Ricci curvature bounds for metric measure spaces introduced in [5]. We will rely on the powerful machinery of calculus on metric measure spaces already developed by Ambrosio, Gigli and Savaré in a series of recent works.

To focus on the new ideas and keep technical difficulties to a minimum we will assume that the metric measure spaces under consideration have unit mass and finite variance. More precisely, we set here

$$\mathcal{X} = \{(X, d, m) \text{ mms} : m \in \mathcal{P}_2(X, d)\}.$$

Definition 3.14. We say that a metric measure space $(X, d, m) \in \mathcal{X}$ satisfies the Riemannian curvature-dimension condition $\text{RCD}(K, N)$ if it satisfies any of the equivalent properties of Theorem 3.15 below.

Theorem 3.15. Let $(X, d, m) \in \mathcal{X}$. The following properties are equivalent:

- (i) (X, d, m) is a strong $\text{CD}^*(K, N)$ space and the semigroup \mathcal{H}_t on $\mathcal{P}_2(X, d)$ is additive.
- (ii) (X, d, m) is a strong $\text{CD}^e(K, N)$ space and the semigroup \mathcal{H}_t on $\mathcal{P}_2(X, d)$ is additive.
- (iii) (X, d, m) is a length space and any $\mu \in \mathcal{P}_2(X, d)$ is the starting point of an $\text{EVI}_{K, N}$ gradient flow of U .

Remark 3.16. Here we mean by additivity that $\mathcal{H}_t(\lambda\mu + (1 - \lambda)\nu) = \lambda\mathcal{H}_t\mu + (1 - \lambda)\mathcal{H}_t\nu$ for any $\mu, \nu \in \mathcal{P}_2(X, d)$ and $\lambda \in [0, 1]$. Since both $\text{CD}^*(K, N)$ and $\text{CD}^e(K, N)$ imply the $\text{CD}(K, \infty)$ condition, Theorem [5, Thm. 6.1] shows that additivity of the semigroup \mathcal{H}_t can equivalently be replaced in (i) and (ii) by the requirement that the Cheeger energy Ch is quadratic.

Proof. (i) \Leftrightarrow (ii): This is a direct consequence of Corollary 3.12.

(ii) \Rightarrow (iii). By Theorem 2.12 it is sufficient to show that $\mathcal{H}_t(\mu)$ is an $\text{EVI}_{K, N}$ gradient flow of U for every $\mu \in \mathcal{P}_2(X, d, m)$ with density uniformly bounded away from 0 and ∞ . Choose $\mu = \rho m \in \mathcal{P}_2(X, d, m)$ with $0 < c \leq \rho \leq C < \infty$ and set $\mu_t := \mathcal{H}_t(\mu)$. By Proposition 2.11 it is sufficient to take reference measures in (2.12) of the form $\sigma = \eta m$ where η is bounded and has bounded support. For any $t > 0$ choose a geodesic $\Gamma^t : [0, 1] \rightarrow \mathcal{P}_2(X, d, m)$ given by the $\text{CD}^e(K, N)$ condition which connects μ_t and σ . Taking into account (2.15) we have to show that for a.e. $t > 0$:

$$\frac{U_N(\sigma)}{U_N(\mu_t)} \leq c_{K/N}(W_2(\mu_t, \sigma)) - \frac{s_{K/N}(W_2(\mu_t, \sigma))}{N \cdot W_2(\mu_t, \sigma)} \frac{d}{dt} \frac{1}{2} W_2(\mu_t, \sigma)^2.$$

By Lemma 3.17 below we have for every t the inequality

$$\frac{U_N(\sigma)}{U_N(\mu_t)} \leq c_{K/N}(W_2(\mu_t, \sigma)) + \frac{s_{K/N}(W_2(\mu_t, \sigma))}{W_2(\mu_t, \sigma) \cdot U_N(\mu_t)} \cdot \liminf_{s \searrow 0} \frac{U_N(\Gamma_s^t) - U_N(\mu_t)}{s}.$$

Thus the proof is finished if we show that for a.e. $t > 0$:

$$\frac{d}{dt} \frac{1}{2} W_2(\mu_t, \sigma)^2 \leq -N \cdot U_N^{-1}(\mu_t) \cdot \liminf_{s \searrow 0} \frac{U_N(\Gamma_s^t) - U_N(\mu_t)}{s}. \quad (3.16)$$

Under the present assumptions on μ and σ and the $\text{CD}(K, \infty)$ condition (which is weaker than the $\text{CD}^e(K, N)$ condition enforced here by Proposition 3.3) this inequality has basically been proven in [5]. Indeed, [5, Theorem 4.1] yields that for a.e. $t > 0$ and all $\varepsilon > 0$:

$$\frac{d}{dt} \frac{1}{2} W_2(\mu_t, \sigma)^2 \leq \frac{\text{Ch}(\rho_t - \varepsilon \varphi_t) - \text{Ch}(\rho_t)}{\varepsilon}, \quad (3.17)$$

where φ_t is a Kantorovich potential for the optimal transport from μ_t to σ . Moreover, [5, Theorem 4.8] yields that for every $t > 0$ and $\varepsilon > 0$:

$$\liminf_{s \searrow 0} \frac{\text{Ent}(\Gamma_s^t) - \text{Ent}(\mu_t)}{s} \geq \frac{\text{Ch}(\varphi_t) - \text{Ch}(\varphi_t + \varepsilon \rho_t)}{\varepsilon}, \quad (3.18)$$

By K -convexity of Ent along the geodesics Γ^t we have

$$\limsup_{s \searrow 0} \frac{\text{Ent}(\Gamma_s^t) - \text{Ent}(\mu_t)}{s} \leq \text{Ent}(\sigma) - \text{Ent}(\mu_t) - \frac{K}{2} W_2(\mu_t, \sigma)^2.$$

Thus $(\text{Ent}(\Gamma_s^t) - \text{Ent}(\mu_t))^2 = o(s)$ as $s \rightarrow 0$ and a Taylor expansion of $x \mapsto e^{-x/N}$ yields that the right-hand side of (3.16) coincides with the left-hand side of (3.18). To finish the proof we recall that [5, Theorem 5.1] shows that under the $\text{CD}(K, \infty)$ condition additivity of the semigroup \mathcal{H}_t implies that Ch is a quadratic form. Hence the right-hand sides of (3.17) and (3.18) coincide up to $O(\varepsilon)$ as $\varepsilon \rightarrow 0$.

(iii) \Rightarrow (ii). Since by Lemma 2.8 an $\text{EVI}_{K, N}$ flow is in particular an EVI_K flow, [5, Theorem 5.1] already gives additivity of the semigroup \mathcal{H}_t . Let us now show that (X, d, m) is a strong $\text{CD}^e(K, N)$ space. The same argument as in the proof of [5, Lemma 5.2] yields for any pair

$\mu_0, \mu_1 \in D(\text{Ent}) \subset \mathcal{P}_2(X, d, m)$ the existence of a geodesic $\Gamma : [0, 1] \rightarrow D(\text{Ent})$ connecting μ_0 to μ_1 . Hence $D(\text{Ent})$ is a geodesic space and Theorem 2.16 shows that (3.1) holds along any geodesic in $D(\text{Ent})$. \square

Lemma 3.17. *Let (X, d, m) satisfy the $\text{CD}^e(K, N)$ condition and let $\mu_0, \mu_1 \in \mathcal{P}_2(X, d, m)$. Then there exists a geodesic $\Gamma : [0, 1] \rightarrow \mathcal{P}_2(X, d, m)$ connecting μ_0 and μ_1 such that, with $\Theta = W_2(\mu_0, \mu_1)$,*

$$U_N(\mu_1) \leq c_{K/N}(\Theta) \cdot U_N(\mu_0) + \frac{s_{K/N}(\Theta)}{\Theta} \cdot \liminf_{t \searrow 0} \frac{U_N(\Gamma_t) - U_N(\mu_0)}{t}. \quad (3.19)$$

Proof. Let Γ be the geodesic connecting μ_0 and μ_1 given by the $\text{CD}^e(K, N)$ condition. (3.1) immediately yields that for every $t \in [0, 1]$:

$$U_N(\Gamma_t) - U_N(\mu_0) \geq \left[\sigma_{K/N}^{(1-t)}(W) - 1 \right] \cdot U_N(\mu_0) + \sigma_{K/N}^{(t)}(W) \cdot U_N(\mu_1).$$

Dividing by t on both sides and passing to the limit $t \searrow 0$ the assertion follows from the fact that

$$\frac{d}{dt} \sigma_{K/N}^{(t)}(\theta) = -\frac{\theta \cdot c_{K/N}(t\theta)}{s_{K/N}(\theta)}, \quad \sigma_{K/N}^{(0)}(\theta) = 0, \quad \sigma_{K/N}^{(1)}(\theta) = 1.$$

\square

Proposition 3.18 (Weighted spaces). *Let $(X, d, m) \in \mathcal{X}$ be a $\text{RCD}(K, N)$ space and let $V : X \rightarrow \mathbb{R}$ be continuous, bounded below and (K', N') -convex function in the sense of Definition 2.15 with $\int \exp(-V) dm = 1$. Then $(X, d, e^{-V}m)$ is a $\text{RCD}(K + K', N + N')$ space.*

Proof. By Proposition 3.4 $(X, d, e^{-V}m)$ is a strong $\text{CD}^e(K + K', N + N')$ space. Invariance of the weak upper gradient under multiplicative changes of the reference measure by [4, Lem. 4.11] together with the Leibnitz rule (3.14) give that the Cheeger energy associated to $e^{-V}m$ is again quadratic. See also [5, Prop. 6.19]. Thus the assertion follows from Theorem 3.15 (ii). \square

The Riemannian curvature-dimension condition has a number of natural properties that we collect here. The first one is the stability under convergence of metric measure spaces in the transportation distance \mathbb{D} . We refer to [24, Sec. 3] for the definition and properties of the transportation distance.

Theorem 3.19 (Stability). *Let $(X_n, d_n, m_n) \in \mathcal{X}$ be a sequence of $\text{RCD}(K, N)$ spaces. If $\mathbb{D}((X_n, d_n, m_n), (X, d, m)) \rightarrow 0$ for some metric measure space $(X, d, m) \in \mathcal{X}$ then (X, d, m) is also a $\text{RCD}(K, N)$ space.*

Proof. We follow essentially the arguments of Ambrosio, Gigli and Savaré in [5, Thm. 6.10] where stability of the $\text{RCD}(K, \infty)$ condition has been established.

We show stability of characterization (iii) in Theorem 3.15. By Proposition 2.11 and Corollary 2.14 it is sufficient to show that for any $\mu = \rho m \in \mathcal{P}_2(X, d, m)$ with $\rho \in L^\infty(X, m)$ there exist a continuous curve $(\mu_t)_{t \in [0, \infty)}$ in $\mathcal{P}_2(X, d)$, locally absolutely continuous in $(0, \infty)$ and starting in μ such that for any $\nu = \sigma m \in \mathcal{P}_2(X, d)$ with $\sigma \in L^\infty(X, d, m)$ and any $s \leq t$:

$$e_K(t-s) \frac{N}{2} \left(1 - \frac{U_N(\nu)}{U_N(\mu_t)} \right) \geq e^{K(t-s)} s_{K/N} \left(\frac{1}{2} W_2(\mu_t, \nu) \right)^2 - s_{K/N} \left(\frac{1}{2} W_2(\mu_s, \nu) \right)^2. \quad (3.20)$$

Choose optimal couplings (\hat{d}_n, q_n) of (X_n, d_n, m_n) and (X, d, m) . Given $\mu = \rho m \in \mathcal{P}_2(X, d, m)$ we set

$$Q_n \mu(dx) = \int \rho(y) q_n(dx, dy) \in \mathcal{P}_2(X_n, d_n, m_n).$$

Similarly we obtain an operator $Q'_n : \mathcal{P}_2(X_n, d_n, m_n) \rightarrow \mathcal{P}_2(X, d, m)$, see [24, Lem. I4.19] and also [5, Prop. 2.2, 2.3].

Now set $\mu^n = Q_n \mu$. By assumption there exists a curve $(\mu_t^n)_{t \in [0, \infty)}$ in $\mathcal{P}_2(X_n, d_n)$ starting from μ^n such that for all $s \leq t$:

$$e_K(t-s) \frac{N}{2} \left(1 - \frac{U_N^n(\nu^n)}{U_N^n(\mu_t^n)} \right) \geq e^{K(t-s)} s_{K/N} \left(\frac{1}{2} W_2(\mu_t^n, \nu^n) \right)^2 - s_{K/N} \left(\frac{1}{2} W_2(\mu_s^n, \nu^n) \right)^2, \quad (3.21)$$

where $\nu^n = Q_n \nu$ and U_N^n corresponds to the relative entropy functional in (X_n, d_n, m_n) . By the maximum principle we have $\mu_t^n \leq C m_n$. For each $t \geq 0$ set $\tilde{\mu}_t^n := Q'_n \mu_t^n \in \mathcal{P}_2(X, d)$. We claim that, after extraction of a subsequence, we have that $\tilde{\mu}_t^n \rightarrow \mu_t$ in $\mathcal{P}_2(X, d)$ as $n \rightarrow \infty$ for curve (μ_t) in $\mathcal{P}_2(X, d)$.

Indeed, note that $\tilde{\mu}_t^n \leq C m$ for all n and t . From the Energy Dissipation Equality (2.10) we conclude that

$$\int_s^t |\dot{\mu}_r^n|^2 dr \leq U^n(\mu^n) \leq C \log C$$

and hence the curves (μ_t^n) are equi-absolutely continuous. Since $m \in \mathcal{P}_2(X, d)$, the set of measures $\{\mu \in \mathcal{P}_2(X, d, m) : \mu \leq C m\}$ is relatively compact w.r.t W_2 -convergence. Hence, by a diagonal argument, we conclude that up to extraction of a subsequence $\tilde{\mu}_t^n \rightarrow \mu_t$ for all $t \in \mathbb{Q}_+$ and some $\mu_t \in \mathcal{P}_2(X, d)$. Using the equi-absolute continuity of the curves (μ_t^n) and the equi-continuity of the maps Q_n we obtain convergence for all times $t \in [0, \infty)$ for the same subsequence and a curve (μ_t) in $\mathcal{P}_2(X, d)$ which is again absolutely continuous.

Finally, we observe that since the operators Q_n, Q'_n do not increase the entropy we have $U^n(\nu^n) \leq U(\nu)$ and by lower semi-continuity of the entropy also $\text{Ent}(\mu_t) \leq \liminf_n \text{Ent}(\tilde{\mu}_t^n) \leq \liminf_n \text{Ent}(\mu_t^n)$. Moreover, we have $W_2(\mu_t^n, \nu^n) \rightarrow W_2(\mu_t, \nu)$. This allows to pass to the limit in (3.21) to obtain (3.20). \square

Theorem 3.20 (Tensorization). *For $i = 1, 2$ let $(X_i, d_i, m_i) \in \mathcal{X}$ be non-branching metric measure spaces satisfying $\text{RCD}(K, N_i)$. Then the product space $(X_1 \times X_2, d, m_1 \otimes m_2) \in \mathcal{X}$, defined by*

$$d((x, y), (x', y'))^2 = d_1(x, x')^2 + d_2(y, y')^2,$$

is also non-branching and satisfies $\text{RCD}(K, N_1 + N_2)$.

Proof. We use characterization (i) of Theorem 3.15. The tensorization property for non-branching $\text{RCD}(K, \infty)$ spaces, see [5, Thm. 6.11], yields that $(X_1 \times X_2, d, m_1 \otimes m_2)$ is again a non-branching $\text{RCD}(K, \infty)$ space. In particular the semigroup \mathcal{H}_t associated to the product space is again additive. Moreover, the reduced curvature-dimension condition tensorizes, i.e. by [8, Thm. 4.1] $(X_1 \times X_2, d, m_1 \otimes m_2)$ satisfies $\text{CD}^*(K, N_1 + N_2)$. Since it is non-branching it is a strong $\text{CD}^*(K, N_1 + N_2)$ space which proves the claim. \square

We conclude with a globalization property of the $\text{RCD}(K, N)$ condition. We say that $(X, d, m) \in \mathcal{X}$ satisfies the Riemannian curvature-dimension condition *locally* if and only if every point $x \in \text{supp } m$ has a neighborhood M such that with $m_M = m(M)^{-1} m|_M$ the metric measure space (M, d, m_M) is a $\text{RCD}(K, N)$ space.

Theorem 3.21 (Local-Global). *Let $(X, d, m) \in \mathcal{X}$ be non-branching and assume that $\mathcal{P}_2(X, d, m)$ is a geodesic space. Then (X, d, m) satisfies the $\text{RCD}(K, N)$ property globally if and only if it satisfies $\text{RCD}(K, N)$ locally.*

Proof. We use again characterization (i) in Theorem 3.15. By [8, Thm 5.1] (X, d, m) satisfies $\text{CD}^*(K, N)$ globally if and only if it satisfies $\text{CD}^*(K, N)$ locally. The same holds for the strong version of the reduced curvature-dimension condition by the non-branching assumption. The globalization and localization of the fact that the semigroup \mathcal{H}_t is additive, or equivalently that Ch is quadratic, is in particular a consequence of [5, Thms. 6.18, 6.20] where global-/localization of the $\text{RCD}(K, \infty)$ condition are shown. \square

4. EQUIVALENCE OF $\text{CD}^e(K, N)$ AND THE BOCHNER INEQUALITY $\text{BI}(K, N)$

4.1. From $\text{CD}^e(K, N)$ to Bakry-Émery-Wang $\text{BEW}(K, N)$ and $\text{BI}(K, N)$. In this section we study the analytic consequences of the Riemannian curvature-dimension condition. In particular, we show that it implies an L^2 -gradient estimate of Bakry-Émery-Wang type. This in turn allows us to establish the full Bochner inequality.

As an immediate consequence of Definition 3.14 and Theorem 2.12 we obtain the following Wasserstein contraction estimate.

Theorem 4.1 (W_2 -expansion bound). *Let (X, d, m) be a $\text{RCD}(K, N)$ space. For any $\mu, \nu \in \mathcal{P}_2(X, d)$ and $0 < s, t$ we have*

$$\begin{aligned} s_{K/N} \left(\frac{1}{2} W_2(\mathcal{H}_t \mu, \mathcal{H}_s \nu) \right)^2 &\leq e^{-K(s+t)} s_{K/N} \left(\frac{1}{2} W_2(\mu, \nu) \right)^2 \\ &\quad + \frac{N}{K} \left(1 - e^{-K(s+t)} \right) \frac{(\sqrt{t} - \sqrt{s})^2}{2(s+t)}. \end{aligned} \quad (4.1)$$

In particular, in the limit $s \rightarrow t$ and $\nu \rightarrow \mu$ we have

$$\begin{aligned} W_2(\mathcal{H}_t \mu, \mathcal{H}_s \nu)^2 &\leq e^{-2Kt} W_2(\mu, \nu)^2 + \frac{N}{K} \frac{1 - e^{-2Kt}}{4t^2} \cdot |s - t|^2 \\ &\quad + o(W_2(\mu, \nu)^2 + |t - s|^2). \end{aligned} \quad (4.2)$$

Remark 4.2. Note that the (K, N) -contraction estimate (4.1) for fixed K and N implies the following (standard) L^2 -Wasserstein K -contraction:

$$W_2(\mathcal{H}_t \mu, \mathcal{H}_t \nu) \leq e^{-Kt} W_2(\mu, \nu). \quad (4.3)$$

Indeed, (4.1) with $s = t$ yields

$$s_{K/N} \left(\frac{1}{2} W_2(\mathcal{H}_t \mu, \mathcal{H}_t \nu) \right) \leq e^{-Kt} s_{K/N} \left(\frac{1}{2} W_2(\mu, \nu) \right).$$

Let $(\sigma_s)_{s \in [0,1]}$ be a W_2 -geodesic with $\sigma_0 = \mu$ and $\sigma_1 = \nu$. We then apply the last inequality with $\mu = \sigma_{k/n}, \nu = \sigma_{(k+1)/n}$ for $k, n \in \mathbb{N}, 0 \leq k \leq n$ and sum with respect to k . Then we obtain (4.3) by letting $n \rightarrow \infty$ since $s_{K/N}(\theta) = \theta + o(\theta^2)$ as $\theta \rightarrow 0$.

Remark 4.3. By virtue of (4.3), several regularity properties for \mathbb{H}_t and \mathcal{H}_t have been obtained in [5] for $\text{RCD}(K, \infty)$ spaces. When $f \in L^2 \cap L^\infty$, $\mathbb{H}_t f$ has a continuous representative, denoted by $\tilde{\mathbb{H}}_t f$, defined as follows (see [5, Theorem 6.1]):

$$\tilde{\mathbb{H}}_t f(x) := \int_X f \, d\mathcal{H}_t(\delta_x).$$

Moreover, $(t, x) \mapsto \tilde{\mathbb{H}}_t f(x)$ belongs to $C_b((0, \infty) \times X)$ for each $f \in L^\infty$. The (standard) Bakry-Émery gradient estimate [5, Theorem 6.2] holds: for $f \in D(\text{Ch})$,

$$|\nabla \mathbb{H}_t f|_w^2 \leq e^{-2Kt} \mathbb{H}_t (|\nabla f|_w^2). \quad (4.4)$$

m -a.e. According to [5, Thm. 6.8] we even have that $\tilde{\mathbb{H}}_t f$ is Lipschitz for any $f \in L^\infty(X, m)$.

Under the stronger $\text{RCD}(K, N)$ condition these regularity properties can be further improved, namely we have that $\tilde{\mathbb{H}}_t f$ is Lipschitz for any $f \in L^2(X, m)$. This can be seen along the lines indicated in [5, Rem. 6.4]. Namely, using that $\text{RCD}(K, N)$ spaces are doubling (see Proposition 3.7) and support a Poincaré inequality by recent work of Rajala [22], one can show that the semigroup \mathbb{H}_t is regularizing from L^1 to L^∞ .

Theorem 4.4 (L^2 -gradient estimate). *Let (X, d, m) be a $\text{RCD}(K, N)$ space. For any $f \in D(\text{Ch})$ and $t > 0$ we have m -a.e. in X :*

$$|\nabla \mathbb{H}_t f|_w^2 + \frac{4Kt^2}{N(e^{2Kt} - 1)} |\Delta \mathbb{H}_t f|^2 \leq e^{-2Kt} \mathbb{H}_t (|\nabla f|_w^2). \quad (4.5)$$

Proof. We first consider the case that f is bounded and Lipschitz. For $x, y \in X$, $x \neq y$ and $t, s \geq 0$ and any coupling $\pi_{s,t}$ of $\mathcal{H}_s(\delta_x)$ and $\mathcal{H}_t(\delta_y)$, we have

$$\tilde{H}_s f(x) - \tilde{H}_t f(y) \leq \int_{X \times X} |f(z) - f(w)| \pi_{s,t}(dzdw). \quad (4.6)$$

Since $|f(z) - f(w)| \leq \text{Lip}(f)d(x, y)$, (4.6) and (4.1) yield

$$\begin{aligned} s_{K/N} \left(\frac{1}{2\text{Lip}(f)} (\tilde{H}_s f(x) - \tilde{H}_t f(y)) \right)^2 &\leq s_{K/N} \left(\frac{1}{2} W_1(\mathcal{H}_s(\delta_x), \mathcal{H}_t(\delta_y)) \right)^2 \\ &\leq s_{K/N} \left(\frac{1}{2} W_2(\mathcal{H}_t(\delta_x), \mathcal{H}_s(\delta_y)) \right)^2 \\ &\leq e^{-K(s+t)} s_{K/N} \left(\frac{1}{2} d(x, y) \right)^2 + \frac{N(1 - e^{-K(s+t)})}{2K(s+t)} (\sqrt{t} - \sqrt{s})^2. \end{aligned}$$

It implies that the map $u \mapsto \tilde{H}_u f(z)$ is locally Lipschitz on $(0, 1)$ for each $z \in X$ and hence differentiable \mathcal{L}^1 -a.e.

The first step is to show the following inequality:

$$|\nabla \tilde{H}_t f|(x)^2 + \frac{4Kt^2}{N(e^{2Kt} - 1)} \left(\frac{\partial}{\partial t} \tilde{H}_t f(x) \right)^2 \leq e^{-2Kt} \tilde{H}_t(|\nabla f|^2)(x) \quad (4.7)$$

for each $x \in X$ and $t > 0$ such that $u \mapsto \tilde{H}_u f(x)$ is differentiable at t . Let $y \in X$ and $s \geq 0$. let us define $r = r(x, y; s, t) > 0$ and $G_r f : X \rightarrow \mathbb{R}$ by

$$\begin{aligned} r &:= \begin{cases} W_2(\mathcal{H}_s(\delta_x), \mathcal{H}_t(\delta_y))^{1/2} & \text{if } W_2(\mathcal{H}_s(\delta_x), \mathcal{H}_t(\delta_y)) > 0, \\ d(x, y) & \text{otherwise.} \end{cases} \\ G_r f(z) &:= \sup_{z'; d(z, z') \in (0, r)} \frac{|f(z) - f(z')|}{d(z, z')}. \end{aligned}$$

Then by taking a coupling $\pi_{s,t}$ as a minimizer of $W_2(\mathcal{H}_s(\delta_x), \mathcal{H}_t(\delta_y))$ in (4.6),

$$\begin{aligned} &\int_{X \times X} |f(z) - f(w)| \pi_{s,t}(dzdw) \\ &= \int_{X \times X} |f(z) - f(w)| 1_{\{d(z, w) \leq r\}} \pi_{s,t}(dzdw) \\ &\quad + \int_{X \times X} |f(z) - f(w)| 1_{\{d(z, w) > r\}} \pi_{s,t}(dzdw) \\ &\leq \int_{X \times X} G_r f(z) d(z, w) \pi_{s,t}(dzdw) + 2\|f\|_\infty \pi_{s,t}(d > r) \\ &\leq \left(\int_X (G_r f)^2 d\mathcal{H}_s(\delta_x) \right)^{1/2} W_2(\mathcal{H}_t(\delta_x), \mathcal{H}_t(\delta_y)) \\ &\quad + \frac{2\|f\|_\infty}{r^2} W_2(\mathcal{H}_s(\delta_x), \mathcal{H}_t(\delta_y))^2. \end{aligned} \quad (4.8)$$

After substituting (4.8) into (4.6), we apply (4.1) with $\mu = \delta_y$ and $\nu = \delta_x$ to obtain

$$\begin{aligned} &\tilde{H}_s f(x) - \tilde{H}_t f(y) \\ &\leq \tilde{H}_s((G_r f)^2)(x)^{1/2} \\ &\quad \times 2s_{K/N}^{-1} \left(\sqrt{e^{-K(s+t)} s_{K/N} \left(\frac{1}{2} d(x, y) \right)^2 + \frac{N(1 - e^{-K(s+t)})}{2K(s+t)} (\sqrt{t} - \sqrt{s})^2} \right) \\ &\quad + 2\|f\|_\infty W_2(\mathcal{H}_s(\delta_x), \mathcal{H}_t(\delta_y)) \end{aligned} \quad (4.9)$$

by using our choice of r .

Since the inequality (4.7) is quadratic w.r.t. scalar multiplication of f , we may assume without loss of generality that

$$|\nabla f|(x) = \limsup_{y \rightarrow x} \frac{[f(x) - f(y)]_+}{d(x, y)}.$$

Take a sequence $(y_n)_{n \in \mathbb{N}}$ in X such that $\lim_{n \rightarrow \infty} \frac{f(x) - f(y_n)}{d(x, y_n)} = |\nabla f|(x)$ holds. Take $\alpha \in \mathbb{R} \setminus \{0\}$, which will be specified later. For each $n \in \mathbb{N}$, let us take $s_n = t + \alpha d(x, y_n)$ and $r_n = r(x, y_n; s_n, t)$. Then we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\tilde{H}_{s_n} f(x) - \tilde{H}_t f(y_n)}{d(x, y_n)} &= \lim_{n \rightarrow \infty} \left(\alpha \frac{\tilde{H}_{s_n} f(x) - \tilde{H}_t f(x)}{s_n - t} + \frac{\tilde{H}_t f(x) - \tilde{H}_t f(y_n)}{d(x, y_n)} \right) \\ &= \alpha \frac{\partial}{\partial t} \tilde{H}_t f(x) + |\nabla \tilde{H}_t f|(x). \end{aligned}$$

Take $\varepsilon > 0$ arbitrary. Since $G_r f$ is non-decreasing in r , by substituting $s = s_n$, $y = y_n$ into (4.9), dividing both sides by $d(x, y_n)$ and letting $n \rightarrow \infty$, we obtain

$$\begin{aligned} \alpha \frac{\partial}{\partial t} \tilde{H}_t f(x) + |\nabla \tilde{H}_t f|(x) &\leq \tilde{H}_t(|G_\varepsilon f|^2)(x)^{1/2} \\ &\quad \times \sqrt{e^{-2Kt} + \alpha^2 \frac{N(1 - e^{-2Kt})}{4Kt^2}}. \end{aligned}$$

Here we used the fact that $\tilde{H}_u(|G_\varepsilon f|^2)$ is continuous in u . Let v_α be a unit vector in \mathbb{R}^2 of the form $\lambda(1, \alpha \sqrt{N(e^{2Kt} - 1)/(4Kt^2)})$ with $\lambda > 0$. Then, by rewriting the last inequality after $\varepsilon \downarrow 0$, we obtain

$$v_\alpha \cdot \left(|\nabla \tilde{H}_t f|(x), \sqrt{\frac{4Kt}{N(e^{2Kt} - 1)}} \frac{\partial}{\partial t} \tilde{H}_t f(x) \right) \leq e^{-Kt} \tilde{H}_t(|\nabla f|^2)(x)^{1/2}.$$

By optimizing this inequality in α , we obtain (4.7).

The second step is to show the following for any bounded and Lipschitz $f \in D(\text{Ch})$: For each $t > 0$ and m -a.e. $x \in X$,

$$|\nabla \tilde{H}_t f|(x)^2 + \frac{4Kt^2}{N(e^{2Kt} - 1)} |\Delta \tilde{H}_t f(x)|^2 \leq e^{-2Kt} \tilde{H}_t(|\nabla f|^2)(x). \quad (4.10)$$

For each $x \in X$, we already know that $t \mapsto \tilde{H}_t f(x)$ is differentiable for \mathcal{L}^1 -a.e. $t \in [0, \infty)$. Thus the Fubini theorem yields that the set $I \subset (0, \infty)$ given by

$$I := \left\{ t \in (0, \infty) \mid t \mapsto \tilde{H}_t f(x) \text{ is differentiable for } m\text{-a.e. } x \in X \right\}$$

is of full \mathcal{L}^1 -measure. Take $t \in I$. Then we have $\frac{\partial}{\partial t} \tilde{H}_t f(x) = \Delta \tilde{H}_t f(x)$ m -a.e. and hence (4.7) yields (4.10). Thus it suffices to show $I = (0, \infty)$ to prove (4.10). Indeed, for any $t \in (0, \infty)$, there is $s \in I$ with $s < t$. Since $(u, z) \mapsto \tilde{H}_u f(z)$ is locally Lipschitz, the dominated convergence theorem implies

$$\tilde{H}_{t-s} \left(\frac{\partial}{\partial t} \tilde{H}_s f \right)(x) = \tilde{H}_{t-s} \left(\lim_{u \rightarrow 0} \frac{\tilde{H}_{s+u} f - \tilde{H}_s f}{u} \right)(x) = \frac{\partial}{\partial t} \tilde{H}_t f(x)$$

and hence $u \mapsto \tilde{H}_u f(x)$ is differentiable at t for any $x \in X$.

Finally we prove the assertion for $f \in D(\text{Ch})$. Let $f_n \in D(\text{Ch})$ be a sequence of bounded Lipschitz functions on X converging to f in $W^{1,2}$ strongly and $|\nabla f_n| \rightarrow |\nabla f|_w$ in L^2 . Then $\Delta \tilde{H}_t f_n \rightarrow \Delta \tilde{H}_t f$ in L^2 and hence the conclusion follows (cf. [5, Theorem 6.2]). \square

In [5] a Bochner inequality without dimension term has been established on metric measure spaces with Riemannian Ricci curvature bounded below. Using the refined L^2 -gradient estimate of Bakry–Émery–Wang type given by Theorem 4.4 we can improve this result by establishing a *finite dimensional* Bochner inequality in $\text{RCD}(K, N)$ spaces.

Theorem 4.5 (Bochner inequality). *The $\text{RCD}(K, N)$ -condition implies the following Bochner inequality $\text{BI}(K, N)$: for all $f \in D(\Delta)$ with $\Delta f \in W^{1,2}(X, d, m)$ and all $g \in D(\Delta)$ bounded and non-negative with $\Delta g \in L^\infty(X, m)$ we have*

$$\begin{aligned} & \frac{1}{2} \int \Delta g |\nabla f|_w^2 \, dm - \int g \langle \nabla(\Delta f), \nabla f \rangle \, dm \\ & \geq K \int g |\nabla f|_w^2 \, dm + \frac{1}{N} \int g (\Delta f)^2 \, dm. \end{aligned} \quad (4.11)$$

Remark 4.6. Note that the classes of functions f, g appearing in the assumption of the previous theorem are not empty. Namely, they contain all functions of the form $H_t h$ with $h \in L^\infty(X, m)$ and $t > 0$, see Remark 4.3.

Proof. We basically follow the arguments developed in [13] in the setting of Alexandrov spaces.

We will first proof (4.11) for $f \in D(\Delta) \cap C_{\text{Lip}}(X)$ with $\Delta f \in D(\Delta) \cap L^\infty(X, m)$. From (4.5) we obtain immediately

$$\begin{aligned} & \int g |\nabla H_t f|_w^2 \, dm + \frac{4Kt^2}{N(e^{2Kt} - 1)} \int g |\Delta H_t f|^2 \, dm \\ & \leq e^{-2Kt} \int g H_t (|\nabla f|_w^2) \, dm. \end{aligned} \quad (4.12)$$

This will yield (4.11) by subtracting $\int g |\nabla f|_w^2 \, dm$ on both sides, dividing by t and taking the limit $t \searrow 0$. Indeed, we can argue exactly as in the proof of [13, Thm. 4.6], using the Leibnitz rule 3.14 (see [5, Thm. 4.18]), and note in addition that

$$\lim_{t \rightarrow 0} \frac{4Kt}{N(e^{2Kt} - 1)} \int g |\Delta H_t f|^2 \, dm = \frac{2}{N} \int g (\Delta f)^2 \, dm. \quad (4.13)$$

To obtain the estimate (4.11) for general f we can employ the approximation argument from the proof of [13, Thm. 4.6], the additional dimension term posing no difficulty. \square

In the setting of smooth Riemannian manifolds the Bochner inequality is equivalent to a refined Bakry–Émery type gradient estimate due to Wang [26]. Also in our setting we can use the Bochner inequality (4.11) to improve the gradient estimate (4.5). Note that (4.14) is stronger than (4.5) for large t .

Proposition 4.7 (Bakry–Émery–Wang). *For any Riemannian mms (X, d, m) the Bochner inequality $\text{BI}(K, N)$ implies the following Bakry–Émery–Wang gradient estimate $\text{BEW}(K, N)$: for any $f \in D(\text{Ch})$ and $t > 0$ we have m -a.e. in X*

$$|\nabla H_t f|_w^2 + \frac{1 - e^{-2Kt}}{NK} |\Delta H_t f|^2 \leq e^{-2Kt} H_t (|\nabla f|_w^2). \quad (4.14)$$

Proof. We basically follow the semigroup argument of Bakry–Émery and Wang.

Arguing as in the proof of Theorem 4.4 it is sufficient to prove (4.14) for f bounded and Lipschitz. Fix g bounded and non-negative and $\varepsilon > 0$ and consider the function

$$h(s) := e^{-2Ks} \int H_{s+\varepsilon} g |\nabla H_{t-s+\varepsilon} f|^2 \, dm.$$

Note that by Remark 4.6 for any $r > 0$ the functions $H_r g, H_r f$ are admissible in (4.11). Thus we can estimate the derivative of h as:

$$\begin{aligned}
h'(s) &= -2K e^{-2Ks} \int H_{s+\varepsilon} g |\nabla H_{t-s+\varepsilon} f|^2 dm \\
&\quad + e^{-2Ks} \int \Delta H_{s+\varepsilon} g |\nabla H_{t-s+\varepsilon} f|^2 dm \\
&\quad - 2e^{-2Ks} \int H_{s+\varepsilon} g \langle \nabla H_{t-s+\varepsilon} f, \nabla \Delta H_{t-s+\varepsilon} f \rangle dm \\
&\geq \frac{2}{N} e^{-2Ks} \int H_{s+\varepsilon} g (\Delta H_{t-s+\varepsilon} f)^2 dm \\
&\geq \frac{2}{N} e^{-2Ks} \int H_{\varepsilon} g (\Delta H_{t+\varepsilon} f)^2 dm,
\end{aligned}$$

where we have used (4.11) in the first and Jensen's inequality in the second inequality. Integrating from 0 to t we obtain:

$$\begin{aligned}
&\int H_{\varepsilon} g |\nabla H_{t+\varepsilon} f|^2 dm + \frac{1 - e^{-2Kt}}{NK} \int H_{\varepsilon} g (\Delta H_{t+\varepsilon} f)^2 dm \\
&\leq e^{-2Kt} \int H_{t+\varepsilon} g |\nabla H_{\varepsilon} f|^2 dm.
\end{aligned}$$

Finally, letting $\varepsilon \rightarrow 0$ and using that g was arbitrary yields the claim. \square

4.2. From BEW(0, N) to EVI(0, N). In the following two sections, we assume that (X, d, m) is a Riemannian metric measure space (i.e. a mms with linear heat flow) and that the associated heat flow satisfies the Bakry-Émery-Wang gradient estimate BEW(K, N). Since the latter is more restrictive than the classical Bakry-Émery gradient estimate BE(K, ∞), we always can rely on the results in [6]. In particular, we already know that the Riemannian curvature-dimension condition RCD(K, ∞) holds true. Thus Ent is K -convex along each geodesic in $\mathcal{P}_2(X, d)$. To simplify the presentation, we assume $m(X) = 1$. Then $\text{Ent} \geq 0$ and $U_N \leq 1$ on $\mathcal{P}_2(X, d)$. Recall that $U_N(\rho) = \exp(-\frac{1}{N} \text{Ent} \rho)$.

For two reasons, we start with a separate study of the case $K = 0$. Firstly, even under this simplifying assumption the calculations are quite involved and vast. It would be much harder to follow the argumentation if also the curvature effects immediately entered the calculations. Secondly, the case $K = 0$ allows to deduce the $\text{EVI}_{K,N}$ condition 'directly' whereas in the general case, currently we are only able to deduce the 'weaker' $\text{CD}^e(K, N)$ condition. (Due to our full circle of implications, of course, this does not really harm.)

Our approach is strongly inspired by the recent work [6] of Ambrosio, Gigli and Savaré. We follow their presentation and adopt to a large extent their notation. The main difference is that our argument now relies on the analysis of the (nonlinear) gradient flow $(\nu_t)_{t \geq 0}$ for the functional $-U_N$ instead of the analysis of the (linear) heat flow which is the gradient flow $(\mu_t)_{t \geq 0}$ for Ent. Both flows are related to each other via time change:

$$\nu_t = \mu_{T_t}, \quad \partial_t T_t = \frac{1}{N} U_N(\mu_{T_t}).$$

The heat semigroup acting on functions (defined as the gradient flow of the Cheeger energy) as well as the heat semigroup acting on measures (defined as the gradient flow of the Boltzmann entropy) will be denoted by H_t , $t > 0$. The same symbol also is used for the point-wise defined kernel acting on functions.

Given a regular curve $(\rho_s)_{s \in [0,1]} \in \mathcal{P}_2(X, d)$, we define a function ('time change') $\tau : [0, 1] \times [0, \infty) \rightarrow [0, \infty)$ by

$$\partial_t \tau_{s,t} = s \cdot U_N(H_{\tau_{s,t}} \rho_s). \quad (4.15)$$

That is, for each fixed s ,

$$\int_0^{\tau_{s,t}} e^{\frac{1}{N} \text{Ent}(H_r \rho_s)} dr = st.$$

This defines τ uniquely.

Differentiating the latter identity w.r.t. s yields

$$t = e^{\frac{1}{N} \text{Ent}(H_\tau \rho_s)} \dot{t} + \int_0^\tau e^{\frac{1}{N} \text{Ent}(H_r \rho_s)} \cdot \frac{1}{N} \int_X \log H_r f_s H_r \dot{f}_s \, dm \, dr$$

which implies

$$\dot{t} = t \cdot U_N(H_\tau \rho_s) - \frac{1}{N} \int_0^\tau \frac{U_N(H_\tau \rho_s)}{U_N(H_r \rho_s)} \int_X \log H_r f_s H_r \dot{f}_s \, dm \, dr \quad (4.16)$$

Here and in the sequel $f_s = \frac{d\rho_s}{dm}$, $\dot{f}_s = \partial_s f_s$ and $\dot{t}_{s,t} = \partial_s \tau_{s,t}$. Put

$$\rho_{s,\tau} := H_{\tau_{s,t}} \rho_s, \quad f_{s,\tau} = \frac{d\rho_{s,\tau}}{dm} = H_{\tau_{s,t}} f_s.$$

Duality for optimal transports provide estimates for $\frac{1}{2}W^2$ from above and below. On the one hand,

$$\frac{1}{2}W^2(\rho_0, \rho_1) = \frac{1}{2} \inf_\rho \int_0^1 |\dot{\rho}_s|^2 \, ds$$

for absolutely continuous curves $(\rho_s)_{s \in [0,1]}$ connecting ρ_0 and ρ_1 . On the other hand

$$\frac{1}{2}W^2(\rho_0, \rho_{1,\tau}) = \sup_{\varphi_0} \left[\int \varphi_1 d\rho_{1,\tau} - \int \varphi_0 d\rho_0 \right]$$

where $(\varphi_s)_{s \in [0,1]}$ is the Hamilton-Jacobi flow starting in φ_0 .

We use the notion of a regular curve in $\mathcal{P}_2(X, d)$ as introduced in [5, Def. 4.10].

Proposition 4.8. *For each regular curve $(\rho_s)_{s \in [0,1]} \in \mathcal{P}_2(X, d)$ and each $t > 0$*

$$\frac{1}{2}W^2(\rho_0, \rho_{1,\tau}) - \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 \, ds + Nt \cdot [U_N(\rho_0) - U_N(\rho_{1,\tau})] \leq \int_0^1 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 \right] \, ds \quad (4.17)$$

Proof. For all the subsequent calculations, apart from considering regular curves, an additional regularization is requested which truncates the singularities of the logarithm appearing in the definition of the entropy and the Fisher information. To simplify our presentation, we skip the details of this regularization which is performed exactly in the same manner as in [6, Lemma 4.15, Thm. 4.16].

We use the abbreviations $\alpha = \alpha_{s,\tau} = \int_0^1 |\nabla \log f_{s,\tau}|^2 f_{s,\tau} \, dm$ and $\beta = \beta_{s,\tau} = \int \Delta \varphi_s f_{s,\tau} \, dm$. Moreover, we put $\alpha_r = \alpha_{s,r} = \int_0^1 |\nabla \log f_{s,r}|^2 f_{s,r} \, dm$ and $u_r = U_N(\rho_{s,r})$, $u_\tau = U_N(\rho_{s,\tau})$.

Given any (sufficiently regular) Hamilton-Jacobi flow $(\varphi_s)_{s \in [0,1]}$ (which finally will be optimized to yield the optimal transport from ρ_0 to $\rho_{1,\tau}$), the basic inequality $\partial_s \varphi_s \leq -\frac{1}{2}|\nabla \varphi_s|^2$ implies

$$\begin{aligned} (A) &:= \int \varphi_1 d\rho_{1,\tau} - \int \varphi_0 d\rho_{0,\tau} - \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 \, ds \\ &= \int_0^1 \left[-\frac{1}{2} |\dot{\rho}_s|^2 + \int \partial_s (\varphi_s f_{s,\tau}) \, dm \right] \, ds \\ &\leq \int_0^1 \left[-\frac{1}{2} |\dot{\rho}_s|^2 + \int \left(-\frac{1}{2} |\nabla \varphi_s|^2 f_{s,\tau} + H_\tau \varphi_s \dot{f}_s + \dot{t} \Delta H_\tau f_s \cdot \varphi_s \right) \, dm \right] \, ds \\ &= \int_0^1 \left[-\frac{1}{2} |\dot{\rho}_s|^2 - \frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} \, dm \right. \\ &\quad \left. + \int \dot{f}_s \cdot H_\tau \varphi_s \, dm + \beta t u_\tau - \beta \frac{1}{N} \int_0^\tau \frac{u_\tau}{u_r} \int \dot{f}_s \cdot H_r \log f_{s,r} \, dm \, dr \right] \, ds. \end{aligned}$$

Moreover,

$$\begin{aligned}
(B) &:= Nt \cdot [U_N(\rho_0) - U_N(\rho_{1,\tau})] = t \int_0^1 U_N(\rho_{s,\tau}) \partial_s \text{Ent}(\rho_{s,\tau}) ds \\
&\leq t \int_0^1 U_N(\rho_{s,\tau}) \cdot \int \log f_{s,\tau} \cdot [H_\tau \dot{f}_s + \dot{\tau} \Delta H_\tau f_s] dm ds \\
&= \int_0^1 \left[tu_\tau \cdot \int \dot{f}_s \cdot H_\tau \log f_{s,\tau} dm - t^2 u_\tau^2 \alpha \right. \\
&\quad \left. + tu_\tau \alpha \frac{1}{N} \int_0^\tau \frac{u_\tau}{u_r} \int \dot{f}_s \cdot H_r \log f_{s,r} dm dr \right] ds.
\end{aligned}$$

Adding up

$$\begin{aligned}
(A) + (B) &\leq \int_0^1 \left[-\frac{1}{2} |\dot{\rho}_s|^2 - \frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm + tu_\tau (\beta - tu_\tau \alpha) \right. \\
&\quad \left. + \frac{1}{\tau} \int_0^\tau \int \dot{f}_s \cdot \left[H_\tau (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} H_r \log f_{s,r} \right] dm dr \right] ds \\
&\leq \int_0^1 \left[-\frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm + tu_\tau (\beta - tu_\tau \alpha) \right. \\
&\quad \left. + \frac{1}{\tau} \int_0^\tau \frac{1}{2} \int \left| \nabla \left[H_\tau (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} H_r \log f_{s,r} \right] \right|^2 f_s dm dr \right] ds \\
&\leq \int_0^1 \left[-\frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm + tu_\tau (\beta - tu_\tau \alpha) \right. \\
&\quad \left. + \frac{1}{\tau} \int_0^\tau \frac{1}{2} \int \left| \nabla \left[H_{\tau-r} (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} \log f_{s,r} \right] \right|^2 f_{s,r} dm dr \right] ds \\
&\quad - \frac{1}{\tau} \int_0^\tau \frac{r}{N} \int \left| \Delta \left[H_\tau (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} H_r \log f_{s,r} \right] \right|^2 f_s dm dr \right] ds \\
&=: (C_1) + (C_2) + ([D + E]^2) + (F)
\end{aligned}$$

by application of the Bakry-Émery-Wang estimate BEW(0, N) to the kernel H_r . The last term will be estimated as follows

$$\begin{aligned}
(F) &\leq \int_0^1 \left[-\frac{1}{\tau} \int_0^\tau \frac{r}{N} \left| \int \Delta \left[H_\tau (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} H_r \log f_{s,r} \right] f_s dm \right|^2 dr \right] ds \\
&= \int_0^1 \left[-\frac{1}{\tau} \int_0^\tau \frac{r}{N} \left| -\beta + tu_\tau \alpha - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} \alpha_r \right|^2 dr \right] ds \\
&= \int_0^1 \left[-\frac{1}{\tau} \int_0^\tau \frac{r}{N} |\beta - tu_\tau \alpha|^2 \cdot \left| 1 + \frac{\tau}{N} \frac{u_\tau}{u_r} \alpha_r \right|^2 dr \right] ds.
\end{aligned}$$

The second last term $([D + E]^2)$ can be decomposed into

$$(E^2) = \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{1}{2} \frac{\tau^2}{N^2} \left(\frac{u_\tau}{u_r} \right)^2 \alpha_r \cdot (\beta - tu_\tau \alpha)^2 dr \right] ds,$$

$$(2DE) = \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{\tau}{N} \frac{u_\tau}{u_r} (\beta - tu_\tau \alpha)^2 dr \right] ds$$

and

$$\begin{aligned}
(D^2) &= \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{1}{2} \int |\nabla H_{\tau-r}(\varphi_s + tu_\tau \log f_{s,\tau})|^2 f_{s,r} dm dr \right] ds \\
&\leq \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{1}{2} \int |\nabla(\varphi_s + tu_\tau \log f_{s,\tau})|^2 f_{s,\tau} dm dr \right. \\
&\quad \left. - \frac{1}{\tau} \int_0^\tau \frac{\tau-r}{N} \int |\Delta H_{\tau-r}(\varphi_s + tu_\tau \log f_{s,\tau})|^2 f_{s,r} dm dr \right] ds \\
&\leq \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{1}{2} \int |\nabla(\varphi_s + tu_\tau \log f_{s,\tau})|^2 f_{s,\tau} dm dr \right. \\
&\quad \left. - \frac{1}{\tau} \int_0^\tau \frac{\tau-r}{N} \left| \int \Delta H_{\tau-r}(\varphi_s + tu_\tau \log f_{s,\tau}) f_{s,r} dm \right|^2 dr \right] ds \\
&= \int_0^1 \left[\frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm - tu_\tau \beta + \frac{1}{2} t^2 u_\tau^2 \alpha - \frac{1}{\tau} \int_0^\tau \frac{\tau-r}{N} (\beta - tu_\tau \alpha)^2 dr \right] ds
\end{aligned}$$

where we applied again the Bakry-Émery-Wang estimate $\text{BEW}(0, N)$, now to the kernel $H_{\tau-r}$. Summing up everything yields

$$(A) + (B) \leq \int_0^1 \left[-\frac{1}{2} t^2 u_\tau^2 \alpha + \frac{1}{N} (\beta - tu_\tau \alpha)^2 \cdot (G) \right] ds$$

where

$$\begin{aligned}
(G) &:= \int_0^\tau \left[-\frac{r}{\tau} \left(1 + \frac{\tau}{N} \frac{u_\tau}{u_r} \alpha_r \right)^2 + \frac{\tau}{2N} \left(\frac{u_\tau}{u_r} \right)^2 \alpha_r + \frac{u_\tau}{u_r} - \frac{\tau-r}{\tau} \right] dr \\
&= \int_0^\tau \left[-r\tau \left(\frac{1}{N} \frac{u_\tau}{u_r} \alpha_r \right)^2 + \frac{\tau}{2N} \left(\frac{u_\tau}{u_r} \right)^2 \alpha_r - \frac{r}{N} \frac{u_\tau}{u_r} \alpha_r \right] dr \\
&\leq \frac{\tau}{4} \left[\left(\frac{u_\tau}{u_0} \right)^2 - 1 \right].
\end{aligned}$$

For the latter note that $\partial_r \frac{1}{u_r} = -\frac{1}{N u_r} \alpha_r$ and thus

$$-\int_0^\tau \frac{r}{N} \frac{u_\tau}{u_r} \alpha_r dr = \tau - \int_0^\tau \frac{u_\tau}{u_r} dr$$

and

$$\frac{1}{N} \int_0^\tau \left(\frac{u_\tau}{u_r} \right)^2 \alpha_r dr = \frac{1}{2} \left[\left(\frac{u_\tau}{u_0} \right)^2 - 1 \right].$$

That is,

$$\int \varphi_1 d\rho_{1,\tau} - \int \varphi_0 d\rho_{0,\tau} - \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 ds + Nt \cdot [U_N(\rho_0) - U_N(\rho_{1,\tau})] \leq \int_0^1 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 \right] ds.$$

Optimizing w.r.t. φ_0 finally yields the claim. \square

Theorem 4.9. *The Bakry-Émery-Wang estimate $\text{BEW}(0, N)$ implies the $\text{EVI}(0, N)$ inequality: For every $\rho_1 \in \mathcal{P}_2(X, d)$ and all $\rho_0 \in \mathcal{P}_2(X, d)$ we have*

$$\frac{d}{dt} \frac{1}{2} W^2(\rho_0, H_t \rho_1) \Big|_{t=0} \leq N \cdot \left[1 - \frac{U_N(\rho_0)}{U_N(\rho_1)} \right].$$

Proof. Since we already know that the Boltzmann entropy Ent is K -convex along each geodesic in $\mathcal{P}_2(X, d)$, we may require that Ent is bounded and continuous along each curve under consideration. Given any AC^2 -curve $(\rho_s)_{s \in [0,1]}$ in $\mathcal{P}_2(X, d)$ along which Ent is continuous, we approximate it by regular curves $(\rho_s^n)_{s \in [0,1]}$ as in [6, Prop. 4.11] with the additional constraint

$$\text{Ent}(H_t \rho_s^n) \rightarrow \text{Ent}(H_t \rho_s) \quad (4.18)$$

for all s, t as $n \rightarrow \infty$.

Then for each s, t the numbers $\tau_{s,t}^n$ (defined in terms of the curve $(\rho_s^n)_{s \in [0,1]}$) converge to the number $\tau_{s,t}$ (defined in terms of the curve $(\rho_s)_{s \in [0,1]}$).

Given $t > 0$, the estimate from the previous proposition holds true for each of the regular curves $(\rho_s^n)_{s \in [0,1]}$. As $n \rightarrow \infty$, by continuity the same estimate holds true for the curve $(\rho_s)_{s \in [0,1]}$. Thus we are allowed to choose $(\rho_s)_{s \in [0,1]}$ as the geodesic connecting ρ_0 and ρ_1 which yields (after dividing by t)

$$\frac{1}{t} \left[\frac{1}{2} W^2(\rho_0, \rho_{1,\tau}) - \frac{1}{2} W^2(\rho_0, \rho_1) \right] + N \cdot [U_N(\rho_0) - U_N(\rho_{1,\tau})] \leq \int_0^1 \frac{\tau}{4t} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 \right] ds. \quad (4.19)$$

Since always $\tau_{s,t} \leq st$, the above RHS can be estimated from above by

$$\int_0^1 \frac{s}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 \right] ds$$

which converges to 0 as $t \rightarrow 0$ according to dominated convergence (w.r.t. s) and the fact that the Boltzmann entropy is continuous along the heat flow: $t \mapsto \text{Ent}(\rho_{s,t})$ is decreasing and lower semi-continuous. Thus

$$\frac{d}{dt} \frac{1}{2} W^2(\rho_0, H_t \rho_1) \Big|_{t=0} \cdot \frac{d}{dt} \tau_{1,t} \Big|_{t=0} = \frac{d}{dt} \frac{1}{2} W^2(\rho_0, H_{\tau_{1,t}} \rho_1) \Big|_{t=0} \leq N \cdot [U_N(\rho_1) - U_N(\rho_0)].$$

Since $\frac{d}{dt} \tau_{1,t} \Big|_{t=0} = U_N(\rho_1)$, this finally yields the $\text{EVI}_{0,N}$ -inequality

$$\frac{d}{dt} \frac{1}{2} W^2(\rho_0, H_t \rho_1) \Big|_{t=0} \leq N \cdot \left[1 - \frac{U_N(\rho_0)}{U_N(\rho_1)} \right].$$

□

4.3. From $\text{BEW}(K, N)$ to $\text{CD}^e(K, N)$. Throughout this section, (X, d, m) will be a Riemannian mms which satisfies the Bakry-Émery-Wang gradient estimate $\text{BEW}(K, N)$. It implies the (weaker) estimate

$$\frac{1}{2} |\nabla H_r f|^2 + \frac{r}{N} e^{-Kr} H_r |\Delta H_r f|^2 \leq \frac{1}{2} e^{-2Kr} H_r |\nabla f|^2. \quad (4.20)$$

Given a regular curve $(\rho_s)_{s \in [0,1]} \in \mathcal{P}_2(X, d)$, we define as in the previous section the time change $\tau : [0, 1] \times [0, \infty) \rightarrow [0, \infty)$ by

$$\partial_t \tau_{s,t} = s \cdot U_N(H_{\tau_{s,t}} \rho_s).$$

Instead of the classical L^2 -Wasserstein distance

$$\frac{1}{2} W^2(\rho_0, \rho_1) = \inf_{\rho} \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 ds$$

we now will be concerned with the weighted distance

$$\inf_{\rho} \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 \cdot \exp(-2K \cdot \tau_{s,t}) ds.$$

Proposition 4.10. *For each regular curve $(\rho_s)_{s \in [0,1]} \in \mathcal{P}_2(X, d)$ and each $t > 0$*

$$\frac{1}{2} W^2(\rho_0, \rho_{1,\tau}) - \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 \cdot e^{-K\tau} ds + Nt \cdot [U_N(\rho_0) - U_N(\rho_{1,\tau})] \leq \int_0^1 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 + 4\tau \right] \cdot e^{2|K|\tau} ds \quad (4.21)$$

Proof. The proof is completely analogous to the proof of the previous Proposition 3.8. We use the same symbols and indicate the difference. Put

$$(\tilde{A}) := \int \varphi_1 d\rho_{1,\tau} - \int \varphi_0 d\rho_{0,\tau} - \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 \cdot e^{-2K \tau_{s,t}} ds$$

and (B) as before. Then

$$\begin{aligned}
(\tilde{A}) + (B) &\leq \int_0^1 \left[-\frac{1}{2} |\dot{\rho}_s|^2 e^{-2K\tau} - \frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm + tu_\tau (\beta - tu_\tau \alpha) \right. \\
&\quad \left. + \frac{1}{\tau} \int_0^\tau \int \dot{f}_s e^{-K\tau} \cdot \left[H_\tau (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} H_r \log f_{s,r} \right] dm e^{+K\tau} dr \right] ds \\
&\leq \int_0^1 \left[-\frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm + tu_\tau (\beta - tu_\tau \alpha) \right. \\
&\quad \left. + \frac{1}{\tau} \int_0^\tau \frac{1}{2} \int \left| \nabla \left[H_\tau (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} H_r \log f_{s,r} \right] \right|^2 f_s dm e^{+2K\tau} dr \right] ds \\
&\leq \int_0^1 \left[-\frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm + tu_\tau (\beta - tu_\tau \alpha) \right. \\
&\quad \left. + \frac{1}{\tau} \int_0^\tau \frac{1}{2} \int \left| \nabla \left[H_{\tau-r} (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} \log f_{s,r} \right] \right|^2 f_{s,r} dm e^{+2K(\tau-r)} dr \right] ds \\
&\quad - \frac{1}{\tau} \int_0^\tau \frac{r}{N} \int \left| \Delta \left[H_\tau (\varphi_s + tu_\tau \log f_{s,\tau}) - \frac{\tau}{N} (\beta - tu_\tau \alpha) \frac{u_\tau}{u_r} H_r \log f_{s,r} \right] \right|^2 f_s dm e^{+K(2\tau-r)} dr \right] ds \\
&=: (C_1) + (C_2) + ([D + E]^2) + (F)
\end{aligned}$$

by application of the Bakry-Émery-Wang estimate $\text{BEW}(K, N)$ to the kernel H_r . The last term will be estimated as follows

$$(F) \leq \int_0^1 \left[-\frac{1}{\tau} \int_0^\tau \frac{r}{N} |\beta - tu_\tau \alpha|^2 \cdot \left| 1 + \frac{\tau}{N} \frac{u_\tau}{u_r} \alpha_r \right|^2 e^{+K(2\tau-r)} dr \right] ds.$$

The second last term $([D + E]^2)$ can be decomposed into

$$\begin{aligned}
(E^2) &= \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{1}{2} \frac{\tau^2}{N^2} \left(\frac{u_\tau}{u_r} \right)^2 \alpha_r \cdot (\beta - tu_\tau \alpha)^2 e^{+2K(\tau-r)} dr \right] ds, \\
(2DE) &= \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{\tau}{N} \frac{u_\tau}{u_r} (\beta - tu_\tau \alpha)^2 e^{+2K(\tau-r)} dr \right] ds
\end{aligned}$$

and

$$\begin{aligned}
(D^2) &= \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{1}{2} \int |\nabla H_{\tau-r} (\varphi_s + tu_\tau \log f_{s,\tau})|^2 f_{s,r} dm e^{+2K(\tau-r)} dr \right] ds \\
&\leq \int_0^1 \left[\frac{1}{\tau} \int_0^\tau \frac{1}{2} \int |\nabla (\varphi_s + tu_\tau \log f_{s,\tau})|^2 f_{s,\tau} dm dr \right. \\
&\quad \left. - \frac{1}{\tau} \int_0^\tau \frac{\tau-r}{N} \left| \int \Delta H_{\tau-r} (\varphi_s + tu_\tau \log f_{s,\tau}) f_{s,r} dm \right|^2 e^{+K(\tau-r)} dr \right] ds \\
&= \int_0^1 \left[\frac{1}{2} \int |\nabla \varphi_s|^2 f_{s,\tau} dm - tu_\tau \beta + \frac{1}{2} t^2 u_\tau^2 \alpha - \frac{1}{\tau} \int_0^\tau \frac{\tau-r}{N} (\beta - tu_\tau \alpha)^2 e^{+K(\tau-r)} dr \right] ds
\end{aligned}$$

where we applied again the Bakry-Émery-Wang estimate $\text{BEW}(K, N)$, now to the kernel $H_{\tau-r}$. Summing up everything yields

$$(A) + (B) \leq \int_0^1 \left[-\frac{1}{2} t^2 u_\tau^2 \alpha + \frac{1}{N} (\beta - tu_\tau \alpha)^2 \cdot (G) \right] ds$$

where

$$\begin{aligned}
(G) &:= \int_0^\tau \left[-\frac{r}{\tau} \left(1 + \frac{\tau}{N} \frac{u_\tau}{u_r} \alpha_r \right)^2 e^{K(2\tau-r)} + \frac{\tau}{2N} \left(\frac{u_\tau}{u_r} \right)^2 \alpha_r e^{2K(\tau-r)} \right. \\
&\quad \left. + \frac{u_\tau}{u_r} e^{2K(\tau-r)} - \frac{\tau-r}{\tau} e^{K(\tau-r)} \right] dr \\
&\leq \frac{\tau}{4} \left[\left(\frac{u_\tau}{u_0} \right)^2 - 1 + 4\tau \right] \cdot e^{2|K|\tau}.
\end{aligned}$$

That is,

$$\begin{aligned}
&\int \varphi_1 d\rho_{1,\tau} - \int \varphi_0 d\rho_{0,\tau} - \frac{1}{2} \int_0^1 |\dot{\rho}_s|^2 \cdot e^{-2K\tau} ds + Nt \cdot [U_N(\rho_0) - U_N(\rho_{1,\tau})] \\
&\leq \int_0^1 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 + 4\tau \right] \cdot e^{2|K|\tau} ds.
\end{aligned}$$

Optimizing w.r.t. φ_0 finally yields the claim. \square

Proposition 4.11. *For each geodesic $(\rho_s)_{s \in [0,2]} \in \mathcal{P}_2(X, d)$*

$$U_N(\rho_1) \geq \frac{1}{2} U_N(\rho_0) + \frac{1}{2} U_N(\rho_2) + \frac{K}{N} |\dot{\rho}|^2 \cdot \int_0^2 g(s, 1) U_N(\rho_s) ds \quad (4.22)$$

where $g(s, r) = \frac{1}{2} \min\{s(2-r), r(2-s)\}$ denotes the Green function on the interval $[0, 2]$.

Proof. As already demonstrated in the proof of Theorem 3.9, the geodesic $(\rho_s)_{s \in [0,1]}$ can be approximated by regular curves $(\rho_s^n)_{s \in [0,1]}$ in such a way that the estimate (4.21) – valid for each of the regular approximations – also holds true for the geodesic $(\rho_s)_{s \in [0,1]}$. That is, we obtain

$$\begin{aligned}
\frac{1}{2} W^2(\rho_0, \rho_{1,\tau}) - \frac{1}{2} W^2(\rho_0, \rho_1) \cdot \int_0^1 e^{-2K\tau} ds &\leq -Nt \cdot [U_N(\rho_0) - U_N(\rho_{1,\tau})] \\
&\quad + \int_0^1 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 + 4\tau \right] \cdot e^{2|K|\tau} ds.
\end{aligned}$$

An analogous estimate holds true for the geodesic $(\rho_{2-s})_{s \in [0,1]}$

$$\begin{aligned}
\frac{1}{2} W^2(\rho_2, \rho_{1,\tau}) - \frac{1}{2} W^2(\rho_2, \rho_1) \cdot \int_1^2 e^{-2K\tau} ds &\leq -Nt \cdot [U_N(\rho_2) - U_N(\rho_{1,\tau})] \\
&\quad + \int_1^2 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 + 4\tau \right] \cdot e^{2|K|\tau} ds.
\end{aligned}$$

Moreover, since $(\rho_s)_{s \in [0,2]}$ is a geodesic

$$\frac{1}{2} W^2(\rho_0, \rho_1) + \frac{1}{2} W^2(\rho_2, \rho_1) - \frac{1}{2} W^2(\rho_0, \rho_{1,\tau}) - \frac{1}{2} W^2(\rho_2, \rho_{1,\tau}) \leq 0.$$

Adding up these three inequalities (and dividing by t) yields

$$\begin{aligned}
\frac{1}{8} W^2(\rho_0, \rho_2) \cdot \frac{1}{t} \left[2 - \int_0^1 e^{-2K\tau} ds - \int_1^2 e^{-2K\tau} ds \right] &\leq N \cdot [2U_N(\rho_{1,\tau}) - U_N(\rho_0) - U_N(\rho_2)] \\
&\quad + \frac{1}{t} \int_0^1 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 + 4\tau \right] \cdot e^{2|K|\tau} ds \\
&\quad + \frac{1}{t} \int_1^2 \frac{\tau}{4} \left[\left(\frac{U_N(\rho_{s,\tau})}{U_N(\rho_{s,0})} \right)^2 - 1 + 4\tau \right] \cdot e^{2|K|\tau} ds.
\end{aligned}$$

As $t \rightarrow 0$, the two last terms on the RHS will disappear (same argument as in the proof of Theorem 3.9. Lower semi-continuity of the entropy implies that the first term on the RHS in the limit $t \rightarrow 0$ will be bounded from above by

$$N \cdot [2U_N(\rho_1) - U_N(\rho_0) - U_N(\rho_2)].$$

Finally, by the very definition of τ ,

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} \left[2 - \int_0^1 e^{-2K\tau} ds - \int_1^2 e^{-2K\tau} ds \right] &= -2K \int_0^2 \partial_t \tau_{s,t} ds \\ &= -2K \left[\int_0^1 s U_N(\rho_s) ds + \int_1^2 (2-s) U_N(\rho_s) ds \right] \\ &= -4K \int_0^2 g(s, 1) U_N(\rho_s) ds. \end{aligned}$$

Thus we end up with

$$-\frac{K}{2} W^2(\rho_0, \rho_2) \cdot \int_0^2 g(s, 1) U_N(\rho_s) ds \leq N \cdot [2U_N(\rho_1) - U_N(\rho_0) - U_N(\rho_2)].$$

Since $|\dot{\rho}| = W^2(\rho_0, \rho_2)/2$, this proves the claim. \square

Remark 4.12. A simple rescaling argument yields that for each geodesic $(\rho_s)_{s \in [0,1]} \in \mathcal{P}_2$

$$U_N(\rho_{1/2}) \geq \frac{1}{2} U_N(\rho_0) + \frac{1}{2} U_N(\rho_1) + \frac{K}{N} |\dot{\rho}|^2 \cdot \int_0^1 g\left(s, \frac{1}{2}\right) U_N(\rho_s) ds$$

where $g(s, r) = \min\{s(1-r), r(1-s)\}$ now denotes the Green function on the interval $[0, 1]$. Moreover, a slight modification of the previous argument allows to prove that

$$U_N(\rho_r) \geq (1-r)U_N(\rho_0) + rU_N(\rho_1) + \frac{K}{N} |\dot{\rho}|^2 \cdot \int_0^1 g(s, r) U_N(\rho_s) ds \quad (4.23)$$

for each $r \in [0, 1]$.

Theorem 4.13. *The Bakry-Émery-Wang estimate $\text{BEW}(K, N)$ implies the entropic curvature-dimension condition $\text{CD}^e(K, N)$.*

Proof. We want to prove that along each Wasserstein geodesic $(\rho_s)_{s \in [0,1]}$, the function

$$u : s \mapsto U_N(\rho_s)$$

satisfies

$$u'' \leq \lambda u \quad (4.24)$$

with $\lambda = -\frac{K}{N} |\dot{\rho}|^2$. The differential inequality implies that

$$u(s) \geq \sigma_\lambda^{(1-s)} \cdot u(0) + \sigma_\lambda^{(s)} \cdot u(1) \quad (4.25)$$

(with the usual coefficients $\sigma_\lambda^{(s)}$) as well as

$$u(s) \geq (1-s) \cdot u(0) + s \cdot u(1) - \lambda \cdot \int_0^1 g(s, r) u(r) dr \quad (4.26)$$

with $g(s, r) = \min\{s(1-r), r(1-s)\}$ being the Green function on the interval $[0, 1]$. Conversely, if any of these two conditions holds for all geodesics, (4.24) will follow. Actually, for this conclusion, it suffices to verify the latter conditions for $s = \frac{1}{2}$. This is the content of the previous Proposition. \square

REFERENCES

- [1] L. Ambrosio and N. Gigli. *User's guide to optimal transport theory*.
- [2] L. Ambrosio, N. Gigli, A. Mondino, and T. Rajala. Riemannian Ricci curvature lower bounds in metric measure spaces with σ -finite measure. Preprint. To appear in *Trans. Amer. Math. Soc.*
- [3] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.
- [4] L. Ambrosio, N. Gigli, and G. Savaré. Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Preprint at arXiv:1106.2090*, 2011.
- [5] L. Ambrosio, N. Gigli, and G. Savaré. Metric measure spaces with Riemannian Ricci curvature bounded from below. *Preprint at arXiv:1109.0222*, 2011.
- [6] L. Ambrosio, N. Gigli, and G. Savaré. Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds. *Preprint at arXiv:1209.5786*, 2012.
- [7] K. Bacher and K.-Th. Sturm. Ricci bounds for Euclidean and spherical cones. Preprint at arXiv:1103.0197.
- [8] K. Bacher and K.-Th. Sturm. Localization and tensorization properties of the curvature-dimension condition for metric measure spaces. *J. Funct. Anal.*, 259(1):28–56, 2010.
- [9] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.
- [10] F. Cavalletti and K.-Th. Sturm. Local curvature-dimension condition implies measure-contraction property. *J. Funct. Anal.*, 262(12):5110–5127, 2012.
- [11] S. Daneri and G. Savaré. Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.*, 40(3):1104–1122, 2008.
- [12] N. Gigli. On the heat flow on metric measure spaces: existence, uniqueness and stability. *Calc. Var. Partial Differential Equations*, 39(1-2):101–120, 2010.
- [13] N. Gigli, K. Kuwada, and S.-I. Ohta. Heat flow on Alexandrov spaces. *Preprint at arXiv:1008.1319*, 2010.
- [14] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1):1–17, 1998.
- [15] C. Ketterer. Ricci curvature bounds for warped products. Preprint at arXiv:1209.1325.
- [16] K. Kuwada. Duality on gradient estimates and Wasserstein controls. *Journal of Functional Analysis*, 258(11):3758 – 3774, 2010.
- [17] P. Li and S.-T. Yau. On the parabolic kernel of the Schrödinger operator. *Acta Math.*, 156:153–201, 1986.
- [18] J. Lott and C. Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. Math. (2)*, 169(3):903–991, 2009.
- [19] S.-I. Ohta. Finsler interpolation inequalities. *Calc. Var. Partial Differential Equations*, 36:211–249, 2009.
- [20] S.-I. Ohta and K.-Th. Sturm. Bochner–Weitzenböck formula and Li-Yau estimate on Finsler manifolds. Preprint at arXiv:1105.0983.
- [21] A. Petrunin. Alexandrov meets Lott–Villani–Sturm. *Münster J. Math.*, 4:53–64, 2011.
- [22] T. Rajala. Local Poincaré inequalities from stable curvature conditions on metric spaces. *Calc. Var. Partial Differential Equations*, 44(3-4):477–494, 2012.
- [23] T. Rajala and K.-Th. Sturm. Non-branching geodesics and optimal maps in strong $\text{cd}(k, \infty)$ -spaces. Preprint at arXiv:1207.6754.
- [24] K.-Th. Sturm. On the geometry of metric measure spaces. I and II. *Acta Math.*, 196(1):65–177, 2006.
- [25] C. Villani. *Optimal transport, Old and new*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.
- [26] Feng-Yu Wang. Equivalent semigroup properties for the curvature-dimension condition. *Bull. Sci. Math.*, 135(6-7):803–815, 2011.
- [27] H.-C. Zhang and X.-P. Zhu. Yau's gradient estimates on Alexandrov spaces. *J. Diff. Geom.*, 91(3):445–522, 2012.

UNIVERSITY OF BONN, INSTITUTE FOR APPLIED MATHEMATICS, ENDENICHER ALLEE 60, 53115 BONN, GERMANY

E-mail address: erbar@iam.uni-bonn.de

E-mail address: kuwada.kazumasa@ocha.ac.jp

E-mail address: sturm@uni-bonn.de